

STIC-Biotech/ChemLib

~~103381~~ 592 Text 104576 104577

From: Stucker, Jeffrey
Sent: Tuesday, September 09, 2003 8:39 AM
To: STIC-Biotech/ChemLib
Subject: 09/980777

I need to search the art for resistance conveying mutations in HIV-2 protease, specifically at position 90. Unfortunately, my application, 09/980777, does not give the complete sequence of HIV-2 protease. Can you use a publicly available sequence to run this search? The most desirable (but not only) mutation is L90M.

Thanks,
Jeffrey Stucker
1648
CM1 8E01
703-308-4237
Mailbox: 8E12

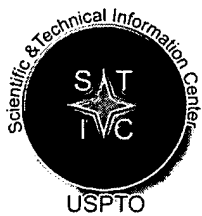
Searcher: D. Schreiber
Phone: 308-4237
Location: CM1 6A03
Date Picked Up: 9/25
Date Completed: 9/25
Searcher Prep/Review: 9/25
Clerical: 9/25
Online time: 19/91

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: 4
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: CompuLink
WWW/Internet: _____
Other (specify): _____

24

THIS PAGE BLANK (USPTO)



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 104576

TO: Jeffrey Stucker
Location: CM1/8E01/8E12
Art Unit: 1648
Friday, September 26, 2003

Case Serial Number: 09/980777

From: David Schreiber
Location: Biotech-Chem Library
CM1-6A03
Phone: 308-4292

david.schreiber@uspto.gov

Search Notes

THIS PAGE BLANK (USPTO)

CC test the ability of various antiviral agents to determined their ability
CC to prevent the virus from fixing on its target. This sequence represents
CC a HIV-2 ROD isolate protein encoded by a polymerase fragment which is
CC described in the method of the invention.
XX
SQ Sequence 1036 AA;

Query Match		99.5%;	Score 5426;	DB 21;	Length 1036;
Best Local Similarity		99.6%;	Pred. No. 0;		
Matches 1032;		Conservative 1;	Mismatches 3;	Indels 0;	Gaps 0;
Qy	1	TGREFRGLKEAPQLPRGPSSAGADTNSPSSSGSGTGEIYAAREKTERAERETIQG	60		
Db	1	TGREFRGLKEAPQLPRGPSSAGADTNSPSSSGSGTGEIYAAREKTERAERETIQG	60		
Qy	61	SDRGLTAPRAGGDTIQGATNGLAAPQFSLWKRPVVTAYIEGQPVVLLDTGADDSIVAG	120		
Db	61	SDRGLTAPRAGGDTIQGATNGLAAPQFSLWKRPVVTAYIEGQPVVLLDTGADDSIVAG	120		
Qy	121	IELGNNYSPKTVGGIGGFINTKEYKNVEIEVLNKKVRATIMTGTPTNIFGRNLTALGM	180		
Db	121	IELGNNYSPKTVGGIGGFINTKEYKNVEIEVLNKKVRATIMTGTPTNIFGRNLTALGM	180		
Qy	181	SLNLPVAKVEPIKIMLPKDGPKLRQWPLTKIEALKEICEKMEKEGQLEEAAPTPNY	240		
Db	181	SLNLPVAKVEPIKIMLPKDGPKLRQWPLTKIEALKEICEKMEKEGQLEEAAPTPNY	240		
Qy	241	NTPTFAKKDKNWRMLIDPRELNKVTQDTEIQLGIPHPAGLAKKRITVLDVGDAF	300		
Db	241	NTPTFAKKDKNWRMLIDPRELNKVTQDTEIQLGIPHPAGLAKKRITVLDVGDAF	300		
Qy	301	SIPLHEDFRPYATFTLPSVNNAPGKRYIYKVLPGQWKGSPAIQHTMRQVLEPFRANK	360		
Db	301	SIPLHEDFRPYATFTLPSVNNAPGKRYIYKVLPGQWKGSPAIQHTMRQVLEPFRANK	360		
Qy	361	DVIIIQYMDITLIASDRDLEHDRVQLKELLNGLGFSTPDERFKQDPPYHWMGYELWP	420		
Db	361	DVIIIQYMDITLIASDRDLEHDRVQLKELLNGLGFSTPDERFKQDPPYHWMGYELWP	420		
Qy	421	TKWLQKIQLPQKEIWTNDIQKLVGLVNNAAQLYPGIKTKHLCLRLIRGKMTLFEVQWT	480		
Db	421	TKWLQKIQLPQKEIWTNDIQKLVGLVNNAAQLYPGIKTKHLCLRLIRGKMTLFEVQWT	480		
Qy	481	ELAEAELEENRIILSQEGHYOEEKELEATVQKQENQWYKIHOEELIKYGYAKV	540		
Db	481	ELAEAELEENRIILSQEGHYOEEKELEATVQKQENQWYKIHOEELIKYGYAKV	540		
Qy	541	KNTHNGIRLLAQVVOQKIGKEALVINGRIPKPHLPVEREIEWQWMDNYQVWTIPDWFV	600		
Db	541	KNTHNGIRLLAQVVOQKIGKEALVINGRIPKPHLPVEREIEWQWMDNYQVWTIPDWFV	600		
Qy	601	STPPLVRLAENLYGDPITPGAETFTYDGSNRSQREGAGYVYDRGDKVKKLEQTTNQQA	660		
Db	601	STPPLVRLAENLYGDPITPGAETFTYDGSNRSQREGAGYVYDRGDKVKKLEQTTNQQA	660		
Qy	661	ELEAFAMALDSDGPKVNIIVDSQVVMGISASQTESSEKIVNQIIEEMIKKEALYVAVWP	720		
Db	661	ELEAFAMALDSDGPKVNIIVDSQVVMGISASQTESSEKIVNQIIEEMIKKEALYVAVWP	720		
Qy	721	AHKGIGGNQEVDLHVSQIGIRQVLFLEKIEPAQEHEHYHSNVKELSHKFGIPNLVAVQIV	780		
Db	721	AHKGIGGNQEVDLHVSQIGIRQVLFLEKIEPAQEHEHYHSNVKELSHKFGIPNLVAVQIV	780		
Qy	781	NSCAQCOQKEAIGHQVNAELGTWQMDCTHLEGIILIVAVHVASGFTAEAVIPQESGRQT	840		
Db	781	NSCAQCOQKEAIGHQVNAELGTWQMDCTHLEGIILIVAVHVASGFTAEAVIPQESGRQT	840		
Qy	841	ALFLLKSLASRPITHLTDCANFTSQEVKVMVNWIGIEQSFYYPNQSGVGYEAMNHH	900		
Db	841	ALFLLKSLASRPITHLTDCANFTSQEVKVMVNWIGIEQSFYYPNQSGVGYEAMNHH	900		
Qy	901	LKNOISRIREQANTTETIVLMAIHCNMFKRGGIGDTPSERLINMITTEQEIQFLOAKN	960		

Db	901	LKNEISRIREQANTTETIVLMAIHCNMFKRGGIGDTPSERLINMITTEQEIQFLOAKN	960	
QY	961	SKLKDFRVYFREGDQLWKGPCELLWKGEAVLVKVGTDIKIIPRRKAKIIRDYGGROEM	1020	
Db	961	SKLKDFRVYFREGDQLWKGPCELLWKGEAVLVKVGTDIKIIPRRKAKIIRDYGGROEM	1020	
QY	1021	DSGSHLEGAREDEGEMA 1036		
Db	1021	DSGSHLEGAREDEGEMA 1036		
RESULT 3				
AAP81773				
ID	AAP81773 standard; protein; 1027 AA.			
XX	AAP81773;			
XX	AAP81773;			
DT	25-MAR-2003 (updated)			
DT	15-NOV-1990 (first entry)			
XX	Sequence encoded by open reading frame of cDNA corresponding to			
DE	HIV-2 ROD genome.			
XX	LAV-II ROD; AIDS; immunogen; antigen; vaccine; diagnostic.			
XX	Human immunodeficiency virus ROD.			
OS	WO8805440-A.			
PN	XX			
PD	28-JUL-1988.			
PF	XX			
PF	15-JAN-1988; 88WO-EP00025.			
XX	XX			
PR	16-JAN-1987; 87US-0003764.			
PR	11-FEB-1987; 87PR-0001739.			
PR	15-APR-1987; 87PR-0005398.			
XX	XX			
PA	(INSP) INST PASTEUR.			
PA	(ALIZ/) ALIZON M.			
PA	(CNRS) CENT NAT RECH SCI.			
XX	XX			
PI	Alizon M, Montagnier L, Guetard D, Clavfl F, Sonigo P, Guyader M;			
PI	Tiollais P, Chakrabarti L, Desrosiers R;			
XX	XX			
DR	WPI; 1988-220290/31.			
DR	N-PSDB; AAN80859.			
XX	XX			
PT	New peptides with immunological properties of HIV-2 envelope protein -			
PT	having the structure of simian immune deficiency virus proteins,			
PT	useful in diagnosis and of vaccine components			
XX	XX			
PS	Disclosure; Fig 1A; 86pp; French.			
XX	XX			
CC	The SQ in AAN80859 was deposited on 21/2/86 at the CNCM under number			
CC	I-522, reference name LAV-II ROD. It is the cDNA to HIV-2 ROD genomic			
CC	RNA. The SQ was compared with the SQ of the genome of SIV (Mac) (AAN80860)			
CC	to identify common regions.			
CC	CC (Updated on 25-MAR-2003 to correct PF field.)			
CC	CC (Updated on 25-MAR-2003 to correct PR field.)			
CC	CC (Updated on 25-MAR-2003 to correct PA field.)			
XX	XX			
SQ	Sequence 1027 AA;			
Query Match 98.7%; Score 5378.5; DB 9; Length 1027;				
Best Local Similarity 98.7%; Pred. No. 0;				
Matches 1023; Conservative 2; Mismatches 2; Indels 9; Gaps 1;				
QY	1	TGREFRTGPGKCAPQLPRGPSSAGADTNSPSSSGSGTGEIYAAREKTERAERETIQG	60	
Db	1	TGREFRTGPGKCAPQLPRGPSSAGADTNSPSSSGSGTGEIYAAREKTERAERETIQG	60	
QY	61	SDRGLTAPRAGGDTIQGATNRGLAAPQFSLWKRPVVTAYIEGQPVVLLDTGADDSIVAG	120	

Dd	61	SDRGLTPRAGGDTIQGATNRGLAAPQFSLWKRPPVVTAYIEGQPVVEVLDTGADDSIVAG	120
Qy	121	IELGNNYSPKIVGGIGGFTNTKEYKNVEVLNKKVRATIMTGTDPINIFGNILTLGGM	180
Dd	121	IELGNNYSPKIVGGIGGFTNTKEYKNVEVLNKKVRATIMTGTDPINIFGNILTLGGM	180
Qy	181	SLNLPVAKVEPIKIMLPGKDGPKLRQWPLTKREKEALKEICEKMEKEGQLEAEPTNPY	240
Dd	181	SLNLPVAKVEPIKIMLPGKDGPKLRQWPLTKREKEALKEICEKMEKEGQLEAEPTNPY	240
Qy	241	NTPPTFAIKKKDNKNWRMLIDFRELKNVTDFTIEIQLGIPHPAGLAKKRITVLDVGDYF	300
Dd	241	NTPPTFAIKKKDNKNWRMLIDFRELKNVTDFTIEIQLGIPHPAGLAKKRITVLDVGDYF	300
Qy	301	SIPLHEDFRPYTAFTLPSVNNAPGKRYIKVLPQGWKGSALFOHTMFOVLEPRKANK	360
Dd	301	SIPLHEDFRPYTAFTLPSVNNAPGKRYIKVLPQGWKGSALFOHTMFOVLEPRKANK	360
Qy	361	DVIIIOYMDIILIASDRTDLEHDRVVLQKELLNGLGFSTPDEKFKQDPYPYHWMGYELWP	420
Dd	361	DVIIIOYMDIILIASDRTDLEHDRVVLQKELLNGLGFSTPDEKFKQDPYPYHWMGYELWP	420
Qy	421	TKWKLOKIQLOPKEIWTNVNDIQKLVGLNWAALPGIKTKHLCLRLIRGKMTLTEEVQWT	480
Dd	421	TKWKLOKIQLOPKEIWTNVNDIQKLVGLNWAALPGIKTKHLCLRLIRGKMTLTEEVQWT	480
Qy	481	ELAEAELENRIILSQEGBHYQOEKELEATVQKDOENQWYKIHQEEKILKVKYAKV	540
Dd	481	ELAEAELENRIILSQEGBHYQOEKELEARVQKDOENWYKIHQEEKILKVKYAKV	540
Qy	541	KNTHNGIRLLAQVQKIGKEALVIGRIPKPHLPVEREIQWQWYVQVWVWIPDWDFV	600
Dd	541	KNTHNGIRLLAQVQKIGKEALVINGRIPKPHLPVEREIQWQWYVQVWVWIPDWDFV	600
Qy	601	STPPLVRLAFNLVGDPIPGAETFTYDGCNRSQKSGKAGYVTRDKDKVKKLEQTTNQOA	660
Dd	601	STPPLVRLAFNLVGDPIPGAETFTYDGCNRSQKSGKAGYVTRDKDKVKKLEQTTNQOA	660
Qy	661	ELEAFAMALTDGPKVNIIVDSQYVNGISASQPTSESKIVNQIIEEMIKKEAIVAWVP	720
Dd	661	ELEAFAMALTDGPKVNIIVDSQYVNGISASQPTSESKIVNQIIEEMIKKEAIVAWVP	720
Qy	721	AHKGGGNEVDHLVSQIRQVLFLEKIEPAQEEHEKYHSNVKELSHKFGIPNLVARQIV	780
Dd	721	AHKGGGNEVDHLVSQIRQVLFLEKIEPAQEEHEKYHSNVKELSHKFGIPNLVARQIV	780
Qy	781	NSCAOQKQGEAIGHGVNAELGTWQMDCTHLEGGKIIIVAVHVASGFIEAEVTPQESGRQT	840
Dd	781	NSCAOQKQGEAIGHGVNAELGTWQMDCTHLEGGKIIIVAVHVASGFIEAEVTPQESGRQT	840
Qy	841	ALFLLLKASRWPTTHLHTONGANFTSOEVKVAWVWIGIEQSGVPYNPQSGQVVGAMNHH	900
Dd	841	ALFLLLKASRWPTTHLHTONGANFTSOEVKVAWVWIGIEQSGVPYNPQSGQVVGAMNHH	900
Qy	901	LKNQISRIREQANTIETIVLMAIHCNFKRRGGIGDMTPSERLINMITTEQIEQIFLOAKN	960
Dd	901	LKNQIS-----ETIVLMAIHCNFKRRGGIGDMTPSERLINMITTEQIEQIFLOAKN	951
Qy	961	SKLKDFRYFRGRQDLWKGPCELLWKGBEGLVVKVGTDIKIIPRKAKIIRDYGGRBQM	1020
Dd	952	SKLKDFRYFRGRQDLWKGPCELLWKGBEGLVVKVGTDIKIIPRKAKIIRDYGGRBQM	1011
Qy	1021	DSGSHLEGAREDEGMA 1036	
Dd	1012	DSGSHLEGAREDEGMA 1027	

RESULT 4
ID AAP80810
XX AAP80810 standard; protein; 1014 AA.
AC AAP80810;
XX

DT	25-MAR-2003	(updated)	
DT	15-NOV-1990	(first entry)	
XX	Sequence of pol protein of HIV-2 ROD isolate CNCM number I-532 in clone HIV-2.P.		
DE			
XX			
KW	AIDS; immunogen; antigen; vaccine; diagnostic.		
XX			
OS	Human immunodeficiency virus ROD.		
XX			
PN	WO8805440-A.		
XX			
PD	28-JUL-1988.		
XX			
PF	15-JAN-1988;	88WO-EP00025.	
XX			
PR	16-JAN-1987;	87US-0003764.	
PR	11-FEB-1987;	87FR-0001739.	
PR	15-APR-1987;	87FR-0005398.	
XX			
PA	(INSP) INST PASTEUR.		
PA	(ALIZ/) ALIZON M.		
PA	(CNRS) CENT NAT RECH SCI.		
XX			
PI	Alizon M, Montagnier L, Guetard D, Clavfl F, Sonigo P, Guyader M; Tiollais P, Chakrabarti L, Desrosiers R;		
PI			
XX	WPI; 1988-220290/31.		
DR			
XX			
PT	New peptide(s) with immunological properties of HIV-2 envelope protein - having the structure of simian immune deficiency virus proteins, useful in diagnosis and of vaccine components		
PT			
PT			
XX			
PS	Disclosure; Fig 5; 86pp; French.		
XX			
CC	The SQ of pol protein of HIV-2 ROD (AAP80810) was compared with the SQ of pol protein of SIV MAC (AAP80809) in Fig 5. New peptides which have immunological properties in common with those of the peptide skeleton of the envelope protein of HIV-2 and also have a peptide structure in common with that of SIV-1 glycoprotein are claimed. Antigenic and immunogenic conjugates contg. the peptides and a kit to detect HIV-2 in biological fluids are new. The peptides are useful for in vitro diagnosis of HIV-2 infection and some of them can be used as components of immunogens and vaccines against HIV. Antibodies raised against them can be used for treatment of AIDS.		
CC	(Updated on 25-MAR-2003 to correct PF field.)		
CC	(Updated on 25-MAR-2003 to correct PR field.)		
CC	(Updated on 25-MAR-2003 to correct PA field.)		
XX			
SQ	Sequence 1014 AA;		
Query Match 97.4%; Score 5310; DB 9; Length 1014;			
Best Local Similarity 99.6%; Pred. No. 0;			
Matches 1010; Conservative 1; Mismatches 3; Indels 0; Gaps 0;			
QY	23	SAGADTNTSPGSSSGSTGEIYAAREKTERAERETIQGSDRGLTAPRAGGDTIQGATNRG	82
Dd	1	SAGADTNTSPGSSSGSTGEIYAAREKTERAERETIQGSDRGLTAPRAGGDTIQGATNRG	60
QY	83	LAAPQFSLWKRPPVVTAYIEGQPVVEVLDTGADDSIVAGIELGNNYSPKIVGGIGGFINTK	142
Dd	61	LAAPQFSLWKRPPVVTAYIEGQPVVEVLDTGADDSIVAGIELGNNYSPKIVGGIGGFINTK	120
QY	143	EYKNVEIEVLNKKVRATIMTGTDPINIFGRNLTALGMSLNLPVAKVEPIKIMLKPGKDG	202
Dd	121	EYKNVEIEVLNKKVRATIMTGTDPINIEFGRNLTALGMSLNLPVAKVEPIKIMLKPGKDG	180
QY	203	PKLRQWPLTKETALKEICEKEKEQLEAEPTNPYNTPTFAIKKKDNKNWRMLIDFR	262
Dd	181	PKLRQWPLTKETALKEICEKEKEQLEAEPTNPYNTPTFAIKKKDNKNWRMLIDFR	240
QY	263	ELNKVQDFTETIQIGIPHPAGLAKKRITVLDVGDAYFSPHEDFRPYTATFLPSVNNNA	322


```
Db 241 ELNKVTDFTQIQLGPHAGLAKRRRTITVLDVGDAYFSIPLHEDFRPYTAFTLPSVNA 300
Qy 323 EPGRRYIKVLPQGWKSPALFOHTMRQVLPFRKANKDVIIIOYMDILLASDRTLEH 382
Db 301 EPGRRYIKVLPQGWKSPALFOHTMRQVLPFRKANKDVIIIOYMDILLASDRTLEH 360
Qy 383 DRVVLQKELLNGLGFTPDEKFKQDPPYHWMGYELWPTKWKLOKIQLPQKEIWTVDIQ 442
Db 361 DRVVLQKELLNGLGFTPDEKFKQDPPYHWMGYELWPTKWKLOKIQLPQKEIWTVDIQ 420
Qy 443 KLGVGLNAAQLYPGIKTKHLCLRLIRKMWLTTEVQWTELAAELEENRIILSQEGHY 502
Db 421 KLGVGLNAAQLYPGIKTKHLCLRLIRKMWLTTEVQWTELAAELEENRIILSQEGHY 480
Qy 503 YOEEKELEATVQKQENQWYKIHQEEKILKVGAKVKNHTNGIRLLAOVQKIGEA 562
Db 481 YOEEKELEATVQKQENQWYKIHQEEKILKVGAKVKNHTNGIRLLAOVQKIGEA 540
Qy 563 LVIWGRIPKPHLPVEREIQWQWYQVWTWIPDWFVSTPPLVRLAFNLVGDPIPAET 622
Db 541 LVIWGRIPKPHLPVEREIQWQWYQVWTWIPDWFVSTPPLVRLAFNLVGDPIPAET 600
Qy 623 FYTDCSNRQSKGAGYVTDGRKDKVKLEQTTNOQAELEAFAMALTDSPKVIIVDS 682
Db 601 FYTDCSNRQSKGAGYVTDGRKDKVKLEQTTNOQAELEAFAMALTDSPKVIIVDS 660
Qy 683 QYVNGISASOPTESKIVNOIIEEMIKKAIYVAVWPAHKGIGGNOEVDHLVSGIRQV 742
Db 661 QYVNGISASOPTESKIVNOIIEEMIKKAIYVAVWPAHKGIGGNOEVDHLVSGIRQV 720
Qy 743 LFLKIEPAQEEHKEYSHNVKELSHKFGIPNLVARIQVNSCAQCOQKGEALHGVNAELG 802
Db 721 LFLKIEPAQEEHKEYSHNVKELSHKFGIPNLVARIQVNSCAQCOQKGEALHGVNAELG 780
Qy 803 TWQMDCTHLGKIIIVAVHVASGFIEAIVPQESGROTALFLKLASRWPTIHLTDNGA 862
Db 781 TWQMDCTHLGKIIIVAVHVASGFIEAIVPQESGROTALFLKLASRWPTIHLTDNGA 840
Qy 863 NTSQEVKVAWWTIGIEQSGVPYNPOSQGVVEAMNHLKNOISRIREQANTIIETVMA 922
Db 841 NTSQEVKVAWWTIGIEQSGVPYNPOSQGVVEAMNHLKNOISRIREQANTIIETVMA 900
Qy 923 IHCNFKRRGGIGDMTPSERLINMITTEQETQFLOAKNSKLKDPVYFREGRODLWKPG 982
Db 901 IHCNFKRRGGIGDMTPSERLINMITTEQETQFLOAKNSKLKDPVYFREGRODLWKPG 960
Qy 983 ELLWKGGAVLVKGTDIKIIPRKAIIIRDYGGREQMDSGSHLEGAREGEMA 1036
Db 961 ELLWKGGAVLVKGTDIKIIPRKAIIIRDYGGREQMDSGSHLEGAREGEMA 1014

RESULT 5
AAP93285
ID AAP93285 standard; protein: 3080 AA.
XX
AC AAP93285;
XX
DT 25-MAR-2003 (updated)
DT 17-DEC-2001 (updated)
DT 06-APR-1990 (first entry)
XX
DE Sequence of clone HIV-2 SBL/ISY.
XX
KW HIV-2; proviral clone HIV-2 SBL/ISY;
XX
OS Human immunodeficiency virus 2.
XX
PN USN7331212-N.
XX
PD 29-AUG-1989.
XX
PF 31-MAR-1989; 89US-0331212.
XX
```

```
PR 30-MAR-1989; 89US-0330446.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICE.
PA (USDC ) US SEC OF COMMERCE.
XX
PI Franchini G, Wongstaal F, Gallo R;
XX
XX WPI: 1989-339698/46.
DR N-PSDB; AAN92119.
XX
XX Complete human immunodeficiency type 2 proviral clone - used to generate
PT animal model for function studies of HIV genes in vivo.
XX
XX Disclosure: Fig. 5; 43pp; English.
XX
XX The protein is encoded by the third reading frame of HIV-2 SBL/ISY, a
CC proviral clone of HIV-2.
CC (Note: Revised entry submitted to correct the patent number format of
CC US Government-owned NTIS applications to prevent clashes with ongoing US
CC granted patent numbers. For further information please visit the Derwent
CC web site at www.derwent.com/dwpi/updates/ntis_us.html.)
CC (Updated on 25-MAR-2003 to correct PA field.)
CC (Updated on 25-MAR-2003 to correct PR field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 3080 AA;

Query Match 93.1%; Score 5075.5; DB 10; Length 3080;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 950; Conservative 49; Mismatches 36; Indels 1; Gaps 1;

Qy 1 TGRFRTGPTLGEAPQLPRGPSSAGADTNSPSSSGSTGEIYAAREKTERAERETIQ 60
Db 584 TGWFFRAWTKGEAPQLPRGPKFAGANTNSPSSSGSTGEVHAAREKTERAETKTQR 643
Qy 61 SDRGLTAPRAGDRTIQGATNRGLAAPQFSLWKRPVVTAYIEGQPEVLLDTGADDSIVAG 120
Db 644 SDRGLAASRRARDTQ-RDDRGLAAPQFSLWKRPVVTAYIEDQPEVLLDTGADDSIVAG 702
Qy 121 IELGNNSPKIVGGIGGINTKEYKNVEIEVLNKKVRATIMTGTPTINIFGRNLTALGM 180
Db 703 IELGNNSPKIVGGIGGINTKEYKDVEIRVLNKKVRATIMTGTPTINIFGRNLTALGM 762
Qy 181 SLNLPVAKVEPIKTLMPGKDGPKLRWPLTKETALKEICEKMEREGQLEEAAPPNTPY 240
Db 763 SLNLPVAKIEPVKTLMPGKDGPKRQWPLFREIEALREICEKMEREGQLEEAAPPNTPY 822
Qy 241 NTPPTFAIKKKDKNKRMLIDFRELNKVTDFTQFTEIQLGIPHAGLAKRRRTITVLDVGDAYF 300
Db 823 NTPPTFAIKKKDKNKRMLIDFRELNKVTDFTQFTEIQLGIPHAGLAKRRRTITVLDVGDAYF 882
Qy 301 SIPLHEDFRPYTAFTLPSVNNAPGKRYIKVLPQGWKGPSAIFQHTMRQVLEPFRKANK 360
Db 883 SIPLYEDFRQYTAFTLPSVNNAPGKRYIKVLPQGWKGPSAIFQYTMROVLEPFRKANP 942
Qy 361 DVIIIOYMDILLASDRTLEHDKVVLQKELLNGLGFTPDEKFKQDPPYHWMGYELWP 420
Db 943 DVIIIOYMDILLASDRTLEHDKVVLQKELLNGLGFTPDEKFKQDPPYHWMGYELWP 1002
Qy 421 TKWKLOKIQLPQKEIWTVDIQLKLVGLNAAQLYPGIKTKHLCLRLIRKMWLTTEVQW 480
Db 1003 TKWKLOKIQLPQKEIWTVDIQLKLVGLNAAQLYPGIKTKHLCLRLIRKMWLTTEVQW 1062
Qy 481 ELAAELEENRIILSQEGHYYOEEKELEATVQKQENQWYKIHQEEKILKVGKAKV 540
Db 1063 ELAAELEENRIILSQEGHYYOEEKELEATVQKQENQWYKIHQEEKILKVGKAKV 1122
Qy 541 KNHTNGIRLLAOVQKIGKALYVGRIPKPHLPVEREIQWQWYQVWTWIPDWFV 600
Db 1123 KNHTNGIRLLAOVQKIGKALYVGRIPKPHLPVEREIQWQWYQVWTWIPDWFV 1182
Qy 601 STPLVRLAFNLVGDPIPGAETFYTDGSCNRSQSKGAGYVTDGRKDKVKLEQTTNOQA 660
Db 1182 STPLVRLAFNLVGDPIPGAETFYTDGSCNRSQSKGAGYVTDGRKDKVKLEQTTNOQA 660
```

1183	STPPLVRLAFNLVKDPIPGAEFTFYDTGSCNRQSKGKAGYITDRGDKVRLIEQTNNQA	1242
661	ELEAFAMALTDSGPKVNIIVDSQVVMGISASQPTSESKIVNQIIEEMIKKEALYVAVWP	720
1243	ELEAFAMAVTDSGPKVNIIVDSQVVMGIVTQPAESERIVNKIIEEMIKKEALYVAVWP	1302
721	AHKGIGNQOEVDHLVSGIRQVLFLEKIEPAQBEHEKYHSNVKLSHKFGIPNLVARQIV	780
1303	AHKGIGNQETDHLVSGIRQVLFLEKIEPAQBEHGKYHSNVKSLAHKFGPLNLVARQIV	1362
781	NSCAQCOCKGEAHGQVNAELGTWQMDCTHLEKIIIVAVHVASGFTAEAVIPOESGRQT	840
1363	NTCAQCOCKGEAHGQVNAELGTWQMDCTHLEKIIIVAVHVASGFTAEAVIPOESGRQT	1422
841	ALFLLKLASRWPITHLHTDNGANFTSQEVKMWAWWVIGIEQSGFYPNPQSQGVVEAMNHH	900
1423	ALFLLKLASRWPITHLHTDNGANFTSQEVKMWAWWVIGIEQSGFYPNPQSQGVVEAMNHH	1482
901	LKQISRIRQANTITETIVLMAHCHNFKRGGIGDMTPSERLINMITTQEITQFOAKN	960
1483	LKQIERIRQANTMETIVLMAVHCHNFKRGGIGDMTPSERLVNMITTQEITQFOAKN	1542
961	SKLDFRVYFREGRDQLWKGPGELLKMGEGAVLVKVGTDIKIIPRRKAKTIIRDYGGROEM	1020
1543	SKLKNFRVYFREGRNQLWQPGELLWKGDGAVIVKVGTDIKVIPRRKAKTIIRDYGGROEM	1602
1021	DSGSHLEGAREDGEMA	1036
1603	DSGSHLEGAREDGEMA	1618

RESULT 6

RESOLUT 0
AAP81771
ID AAP81771 standard; protein; 3210 AA.

XX
AC
AAB91771.

XX 75-MVB-2003 1004-2003

DI	23-MAR-2003	(updated)
DT	07-NOV-1990	(first entry)

DE Deduced sequence encoded by bottom reading frame of cDNA clone
DE HIV-2 SBL/ISY of HIV related retrovirus strain.

KW HIV vaccine; HIV strain SBL-6669-85.

OS Human immunodeficiency virus.

XX
PW
W08808449-A.

XX
PD
03-NOV-1988XX
PF
28-APR-1988: 88WO-SF00218XX
PR 28-APR-1987. 87SE-0001765

XX
PA (SBL-) SBL STATENS BAKTERI.
PA (STAT-) STATENS BAKTERIOLOGISKA LAB.

PI Albert J, Biberfeld G, Fenyo EM, Norrby E;
xx

XX DR WPI; 1988-322769/45.

DR WFL; 1988-322769/
DR N-PSDB; AAN80890;

AA HIV related human retro-virus strain -
PT used for obtaining antigens for assays and vaccines and for
PT prodn. of antibodies for assays
PT

PS Claim 9; Fig 4; 28pp; English.

AA Synthetically produced proteins and peptides, characterised in that the
CC AA sequence is derived from the primary nucleotide sequence of
CC HIV-2 SBL/ISY or a part thereof, or a degenerate thereof are claimed.
CC HIV-2 SBL/ISY represents the complete genome of the virus SBL-6699

(-SBL-6669-85). The proviral DNA was obtd. from a genomic library constructed from DNA of HUT-78 cells infected with SBL-6669-85 using the lambda-phage vector EMBL-3. SBL-6669-85 was isolated from lymphocytes of a West African woman. Protection is requested for the entire genome disclosed in n80890 and for parts thereof, and corresp. to various genes such as the gag gene (corresp. to nucleotides 547 to 2106), the pol gene (nucleotides 1827-4931) and the env gene (nucleotides 5144 to 8682), the corresp. AA sequences and parts thereof and various products derived therefrom, or use thereof, such as clones prepd. by recombinant vector method, HIV test devices and methods. X corresponds to the translation of a stop codon.
CC (Updated on 25-MAR-2003 to correct PA field.)

Sequence	3210 AA;
----------	----------

Query Match 92.8%; Score 5059; DB 9; Length 3210;

Best Local Similarity 91.6%; Pred. No. 0;

Matches	949;	Conservative	49;	Mismatches	36;	Indels	2;	Gaps	2;
---------	------	--------------	-----	------------	-----	--------	----	------	----

Qy	1	TGRFFRTGPI	LKGAEPQLPRGPSSAGADTNPSSGSSSGTGEIYVAAREKTEBRAERETIQ	60
Db	609	TGWFRAWITM	KGKAEPQLPRGPKFAGANTNSTPNGSSSGPTGEVHAAREKTEBRAETKIQ	668
Qy	61	SDRGLTAPRAGGDT	IOGATNRGLAAPQFSLWKRPVVTAITEGQPEVVLDTGADDSIVAG	120
Db	669	SDRGLAASRARDTQ	-RDRGUGLAAPQFSLWKRPVVTAITEQPEVVLDTGADDSIVAG	727
Qy	121	IELGNNTSPKIVGGIGGFINT	KEYKNYVEIEVLNKKVRATIMTGDTPIINFGRNILITALGM	180
Db	728	IELGSNTSPKIVGGIGGFINT	KEYKDYEIRVLNKKVRATIMTGDTPIINFGRNILITALGM	787
Qy	181	SLNLPVAKVBP	IKIMLPGKDGPKLRQWPLITKEIEALKEICEKMEKEGQLEEAAPPNTPY	240
Db	788	SLNLPVAKIEBPVKVITL	KPGDKGPKRQWPLITREKIEALREICEKMEREGQLEEAAPPNTPY	847
Qy	241	NTPTFAIKKDKKNKRWMLID	FRNLNKTQDFTEIOLGPHIPAGLAKKRRITVLDVGDAYF	300
Db	848	NTPTFAIKKDKKNKRWMLID	FRNLNKTQDFTEVOLGPHIPAGLAKKRRITVLDVGDAYF	907
Qy	301	SIPLHEDFRPYTAFTLPSVNA	AEPEGKRYIKYVLPQGWKGSIPAIFQHTMRQVLEPFRKANK	360
Db	908	SIPLYEDFRQYTAFTLPSVNA	AEPEGKRYIKYVLPQGWKGSIPAIFYHTMRQVLEPFRKANP	967
Qy	361	DVIIQYMDILLIASDRDTE	LHRDVVLQKLKELLNGLGSTPDEKFKQDPPYHNMGYELWP	420
Db	968	DVIIQYMDILLIASDRDTE	LHKDKVVLQKLKELLNGLGSTPDEKFKQDPPYQNMGYELWP	1027
Qy	421	TKWKLOKIQIOPKEI	WTVNDIOKLGVGLNWAQAOLYPGIKTKHLCRLLRKMTLFEEOVWT	480
Db	1028	TKWKLOKIQIOPKEI	WTVNDIOKLGVGLNWAQAOLYPGIKTKHLCRLLRKMTPTTEEOVWT	1087
Qy	481	ELAEAELEENRITLSO	EGHYGYOEKLELATVOKDOENQWTKYIHOEEKILKVGKYAKV	540
Db	1088	ELAEAELEENRITLSO	EGHYGYOEKLELATVOKDOENQWTKYVHOGEKILKVGKYAKI	1147
Qy	541	KNPTHNGIRLLAQQVQK	IGKEALVINGRIEKPFLPVEREITWQEWMDNYQVWTIPDWDV	600
Db	1148	KNPTHNGVRLLAQQVQK	IGKEALVINGRIEKPFLPVERETWQEWMDNYQVWTIPDWDV	1207
Qy	601	STPPLVRLAFNLVGDPI	PGAAETFTYDGSCHNRQSKGKAGYVTDGRGDKVKKLEQTTNOQA	660
Db	1208	STPPLVRLAFNLVKDPI	PGAATEFTYDGSCHNRQSKGKAGYITDRGDKVKRILEQTTNOQA	1267
Qy	661	ELEAFAMALTDSPKVINI	YVDSQVYMGISASQPTESKIVNOIIEEMIKKEALYVAVWP	720
Db	1268	ELEAFAMAVTDSGPKVINI	VYDSQVYMGIVTGPAPAESRIVNKRKIEEMIKKEALYVAVWP	1327
Qy	721	AHKGIGGNQVEDHLV	SQIRQVLFLEKIEPAQEEHEKYHSNVAKLSHKFCGIPNLVARQIV	780
Db	1328	AHKGIGGNQIEDHLV	SQIRQVLFLEKIEPAQEEHKGKHSNVAKLSHKFCGIPNLVARQIV	1387
Qy	781	NSCAQCQKQGEALH	GOVNAELGTWQMDCTHLEKIIIVAVHVASGFTAEAVIPQESGRQT	840

Db	1388	NTCAOQCKGEAIGHQVNAELGTQWMDCTHLEGGKIIIVAVHVASGFIEAEVIPSQSGROT	1447
Qy	841	ALFLKLLASRPITHLHTDNGANFTSQEVKVMVAMWIGIEQSPGVYPNQSQGVVAMNHH	900
Db	1448	ALFLKLLASRPITHLHTDNGANFTSQEVKVMVAMWIGIEQSPGVYPNQSQGVVAMNHH	1507
Qy	901	LKNOISRIREQANTTETIVLMAIHCNMFKRGGIGDMPSPERLINMITTEQIEQIOFLQAKN	960
Db	1508	LKNOISRIREQANTTETIVLMAVHCNMFKRGGIGDMPSPERLINMITTEQIEQIOFLQAKN	1567
Qy	961	SKLKDFRVYFREGDRQLWKGPGELLWKGEGAVLVKVGTDIKIIPRRKAKIIRDYGGROEM	1020
Db	1568	SKLKDFRVYFREGDRQLWKGPGELLWKGEGAVLVKVGTDIKIIPRRKAKIIRDYGGROEM	1627
Qy	1021	DSGSHLEGAREGEMA	1036
Db	1628	DSGSHLEGARE-GEMA	1642
RESULT 7			
ID	AAW13055	standard; Protein; 1055 AA.	
XX	AAW13055;		
AC	AAW13055;		
DT	25-MAR-2003 (updated)		
DT	15-OCT-1997 (first entry)		
XX	HIV-2 provirus-encoded pol protein.		
DE	HIV-2 provirus-encoded pol protein.		
XX	HIV-2; HIV-2KR; provirus; gene therapy; vector; vaccine; AIDS;		
KW	pol protein.		
OS	Human immunodeficiency virus type 2.		
XX	WO9705242-A1.		
PN	13-FEB-1997.		
PD	09-JUL-1996; 96WO-US11445.		
XX	07-JUN-1996; 96US-0659251.		
PR	26-JUL-1995; 95US-0001441.		
XX	(REGC) UNIV CALIFORNIA.		
PA			
XX	Kraus G, Wongstaal F, Talbott R, Poeschla E;		
PI			
XX	WPI; 1997-145681/13.		
DR	N-PSDB; AAT61085.		
XX	Human immunodeficiency virus type II proviral clone - useful to		
PT	develop prods. for study, diagnosis, prevention and treatment of HIV		
PT	infections		
PS	Disclosure; Page 89-92; 128pp; English.		
XX	Polypeptides (AAW13052-59) respectively comprise the env, gag, nef,		
CC	pol, rev, tat, vif and vpr polypeptides encoded by HIV-2KR (AAW61085),		
CC	an HIV-2 provirus that comprises a full-length HIV-2 genome.		
CC	HIV-2KR was obtd. from the genomic DNA of Molt 4 Clone 8 lymphoblast		
CC	cells infected with HIV-2P812, an HIV-2 strain isolated from the		
CC	peripheral blood mononuclear cells of an HIV-2 seropositive AIDS		
CC	patient. HIV-2KR polypeptides can be expressed in host cells and		
CC	used as diagnostic reagents and as immunogens for the production of		
CC	diagnostic antibodies.		
CC	(Updated on 25-MAR-2003 to correct PI field.)		
XX			
SQ	Sequence 1055 AA;		

Query Match 92.7%; Score 5055.5; DB 18; Length 1055;
Best Local Similarity 91.5%; Pred. NO. 0;
Matches 948; Conservative 44; Mismatches 43; Indels 1; Gaps 1;

Qy	1	TGRPRTGPTLGRKEAPQLPRGPSSAGADTNSTPSSGSGTGEBIYAAREKTERAERETIQG	60
Db	21	TGWFFRDWPMGKEASQLPRDSPAGADTNSTPSSRPAREVLAAAREAEARENETIQG	80
Qy	61	SDRGUTAPRAGGDTTIGATNRGLAAPQFSLMKRPVVVTAYIEGQPVVEVLDTGADDSIVAG	120
Db	81	DGRGLTAPTRRDTTQRG-DRGFAAPQFSLMKRPVVVTAYVEGQPVVEVLDTGADDSIVAG	139
Qy	121	TELGNYSYSPKIVGGTGGTINTKEYKNVEVLNKKVRATIMTGDTPINIFGRNITLALGM	180
Db	140	TELGSYSPKIVGGTGGTINTKEYKNVEIKVLNKKVKATIMTGDTPINIFGRNITLALGM	199
Qy	181	SUNLPVAKVEPTIKMLKPGKGPURQWPLTKEALKEICEKKEGQLEAPPTNPY	240
Db	200	SUNLPVAKVDPIKVLKPGKGPURQWPLTKEALKEICEKMERGQLEAPPTNPY	259
Qy	241	NTPTFAIKKKDKNKRMLIDFRELNKVTQDFTTEIQLGIPHAGLAKKRITVLDYDAYF	300
Db	260	NTPTFAIKKKDKNKRMLIDFRELNKVTQDFTTEIQLGIPHAGLAKKRITVLDYDAYF	319
Qy	301	SIPLHEDFRPYTAFTLPSVNAEPCKRYIYKVLPOGWKGSIPAFOHTMRQVLEPPRKANK	360
Db	320	SIPLHEDFRQYTAFTLPTVNAEPCKRYIYKVLPOGWKGSIPAFOHTMRQVLEPPRKANK	379
Qy	361	DVIIIOYMDLILIASDRTDLEHDRVVLQKELLNGLGFSTPDEKFKQDPPYHWMGYELWP	420
Db	380	DVILVOYMDLILIASDRTDLEHDRVVLQKELLNGLGFSTPDEKFKQDPPYHWMGYELWP	439
Qy	421	TKWKLQIKLPQKEIWTYNDIOKLVGLNMAAQYLGIKTKHLCLIRLIRGKMTLBEVQMT	480
Db	440	TKWKLQIKLPQKEVMTYNDIOKLVGLNMAAQYLGIKTKHLCLIRLIRGKMTLBEVQMT	499
Qy	481	ELAEAELEENRIILSQEGBGHYQBEKELEATVQKQDENQWYTKIHQSEKILKVGYAKV	540
Db	500	ELAEAELEENRIILSQEGBGHYQBEKELEATVQKQDENQWYTKIHQSEKILKVGYAKV	559
Qy	541	KNHTNGIRLLAQVVKIGKEALVIGRIPFHPVPEREIVEQWMDNWQVWIWDWDFV	600
Db	560	KNHTNGVRLLAHVVKIGKEALVIGRIPFHPVPEREIVEQWMDNWQVWIWDWDFV	619
Qy	601	STPPLVRLAFNLVGDPIPGAETFYDGSNCNQSKEGKAGYVTDGRDKVKVLEQTTNQA	660
Db	620	STPPLVRLAFNLVGDPIPGAETFYDGSNCNQSKEGKAGYVTDGRDKVKVLEQTTNQA	679
Qy	661	ELEAFAMALTDGSGPKVNIIVDSQYVNGISASQPTSESKIVNQIIEEMIKKEAYIVAWVP	720
Db	680	ELEAFAMALTDGSGPKANIIVDSQYVNGIVAGQPTSESKLVNQIIEEMIKKETLYVAVWP	739
Qy	721	AHKGIGGNOEVDHLVSGIROVLFLEKIETPAQEEHEKYHSNVKELSHKFGIPNLVARQIV	780
Db	740	AHKGIGGNOEVDHLVSGIROVLFLEKIETPAQEEHEKYHSNVKELSHKFGIPNLVARQIV	799
Qy	781	NSCAOQCKGEAIGHQVNAELGTWMDCTHLEGGKIIIVAVHVASGFIEAEVIPSQSGROT	840
Db	800	NTCAOQCKGEAIGHQVDAELGTWMDCTHLEGGKIIIVAVHVASGFIEAEVIPSQSGROT	859
Qy	841	ALFLKLLASRPITHLHTDNGANFTSQEVKVMVAMWIGIEQSPGVYPNQSQGVVAMNHH	900
Db	860	ALFLKLLASRPITHLHTDNGANFTSQEVKVMVAMWIGIEQSPGVYPNQSQGVVAMNHH	919
Qy	901	LKNOISRIREQANTTETIVLMAIHCNMFKRGGIGDMPSPERLINMITTEQIEQIOFLQAKN	960
Db	920	LKNOISRIREQANTMETIVLMAVHCNMFKRGGIGDMPSPERLINMITTEQIEQIOFLHAKN	979
Qy	961	SKLKDFRVYFREGDRQLWKGPGELLWKGEGAVLVKVGTDIKIIPRRKAKIIRDYGGROEM	1020
Db	980	SKIKNFRVYFREGDRQLWKGPGELLWKGEGAVLVKVGTDIKIIPRRKAKIIRDYGGRE	1039
Qy	1021	DSGSHLEGAREGEMA	1036
Db	1040	DSSSHLEGTREDGEVA	1055

```

RESULT 8
AAR04025
ID AAR04025 standard; protein; 1035 AA.
XX
AC AAR04025;
XX
DT 25-MAR-2003 (updated)
DT 29-MAY-1989 (first entry)
XX
DE Pol gene product of cDNA to HIV-2 RNA.
XX
KW HIV; AIDS; Vaccine; pUC-HIV-2(GH-1).
XX
OS HIV-2.
XX
PN JP01289486-A.
XX
PD 21-NOV-1989.
XX
PF 16-MAY-1988; 88JP-0119024.
XX
PR 16-MAY-1988; 88JP-0119024.
XX
PA (TOFU ) TOA NENRYO KOGYO KK.
PA (FJRE ) FUJI REBIO KK.
XX
WPI: 1990-005177/01.
DR P-PSDB; AAR04024, AAR04025, AAR04026, AAR04027, AAR04028, AAR04029.
XX
PT DNA indicating complement to RNA gene -
PT of Human Immunodeficiency Virus type 2 used for new vaccine or
PT diagnostic for AIDS virus.
XX
PS Claim 2; Fig 4; 12pp; Japanese.
XX
CC cDNA to novel HIV-2 (GH-1) has been integrated into plasmid
CC pUC HIV-2(GH-1). Useful for diagnosis and vaccination against the
CC virus. Described here is the Pol gene product of the cDNA.
CC See also AAO02830.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 1035 AA;
Query Match 92.1%; Score 5023.5; DB 11; Length 1035;
Best Local Similarity 91.0%; Pred. No. 0;
Matches 944; Conservative 52; Mismatches 38; Indels 3; Gaps 3;

QY 1 TGRFFRTGPGKEAPQLPRGPGSAGADTNSPTSGSSGSTGEIYAAAREKTERAERTIQG 60
DB 1 TGRFFRDGSGKEAPQLPRGPGSSGADTNSPTSRSSSGTGKIYAAAGERAEGAEGTIQR 60
QY 61 SDRGLTAPRAGDGTIOGATNRGLAARQFSLWKRPPVTVATIEGQPVVLLDTGADDSIVAG 120
DB 61 GDGRLTAPRAGKSTSGRQ-DRLAARQFSLWKRPPVTVATIEQPVVLLDTGADDSIVAG 119
QY 121 IEIGNNYSPKIVGGIGFINTKEYKNVEIENKVRATIMTGDTPINIFGRNILTALGM 180
DB 120 IQLGDNYPKIVGGIGFINTKEIKNLKVRATIMTGDTPINIFGRNILTALGM 179
QY 181 SLNLPVAKVEPTKIMLKPGDKPKLRQWPLTKEIEALKEICEKMEKEGQLEAPPTNY 240
DB 180 SLNLPITAKIEPTKIVTLKPGDKGPRLRQWPLTKEIEALKEICEKMEKEGQLEAPPTNY 239
QY 241 NTPTEFAIKKKDKNWRMLDFELNKNVTDQFTEIQIGIPHPAGLAKKRITVLDVGDYF 300
DB 240 NTPTEFAIKKKDKNWRMLDFELNKNVTDQFTEIQIGIPHPAGLAKKRITVLDVGDYF 299
QY 301 SIPLHEDFRPYTAFTLPSVNNAPGKRYIKVLPQGWKGSIPAIFQHTMRQVLEPFRKANK 360
DB 300 SIPLHEDFRQYTAFTLPSVNNAPGKRYIKVLPQGWKGSIPAIFQHTMRQVLEPFRKANK 359
QY 361 DVIIIQYMDILLIASDRTGLEHDKVYVQLKELLNGLGFSTPDEKFKQDPPLQMMGYELWP 420

```

```

DB 360 DVIIIQYMDILLIASDRTGLEHDKVYVQLKELLNGLGFSTPDEKFKQDPPLQMMGYELWP 419
QY 421 TKWKLOKIQLOPOKEITWTVNDIOKLVGLNWAALPYGKTKHLCLRLIRCKMTLTTEVQWT 480
DB 420 TKWKLOKIQLOPOKEITWTVNDIOKLVGLNWAALPYGKTKHLCLRLIRCKMTLTTEVQWT 479
QY 481 ELAEAELEENRIILSQEQEGHYQEEKELEATVQKDQENQWYKIHQEEKILKVGKYAKV 540
DB 480 ELAEAELEENRIILSQEQEGHYQEEKELEATVQKDQENQWYKIHQEEKILKVGKYAKI 539
QY 541 KNHTNGIRLLAQVOKIGKEALVINGRIPKPHLPVEREINWQMDNYQVWTWIPDQWDF 600
DB 540 KNHTNGVRLLAQVQKIGKEALVINGRIPKPHLPVERETWQMDNYQVWTWIPDQWDF 599
QY 601 STPPLVRLAFNLVGDPIPGAETFTYDGSNRSKSGKAGYVTDGRGDKVKKLEOTTNQQA 660
DB 600 STPPLVRLFTNLVGDPIPGAETFTYDGSNRSKSGKAGYVTDGRGDKVRLERTNQQA 659
QY 661 ELEAFAMALTDGPKVNIIVDSQYVMGISASQPTSESKIVNQIIEEMIKKEAIYVAVWP 720
DB 660 ELEAFAMALTDGPKVNIIVDSQYVMGISASQPTSESKIVNQIIEEMIKKEAVYVAVWP 719
QY 721 AHKGIGGQEVNHLVSGIRQVLFLEKIEPAQEEHEKYHSNVKLSHKGEGINLVARQIV 780
DB 720 AHKGIGGQEVNHLVSGIRQVLFLEKIEPAQEEHEKYHSNVKLSHKGEGIPQLVARQIV 779
QY 781 NSCAOQOQGEAIGHGVNAELGTWQMDCTHLEKIIIVAVHVASGFIEAEVPOESGROT 840
DB 780 NTCACQOQGEAIGHGVNAELGTWQMDCTHLEKIIIVAVHVASGFIEAEVPOESGROT 839
QY 841 ALFLKLASRPWITHTLHTDNGANFTSQEVKVAWWTIGIEQSGFVYPNPOSGVVEAMNHH 900
DB 840 ALFLKLASRPWITHTLHTDNGSFTSQEVKVAWWTIGIEQSGFVYPNPOSGVVEAMNHH 899
QY 901 LKNQISRIREQANTTETIVLMAIHCNFKRRGGIDMTSPERLINMITTEQEIQFLOAKN 960
DB 900 LKNQISRIREQANTTETIVLMAIHCNFKRRGGIDMTSPERLINMITTEQEIQFLOAKN 959
QY 961 SKLKDFRVYFREGDQLWKGELWKGAVLVKVGTDIKIIPRKAIIIRDYGGROEM 1020
DB 960 SNFKNFQVYFREGDQLWKGELWKGAVLVKVGTDIKIIPRKAIIIRDYGGROEL 1019
QY 1021 DSGSHLEGAR-EDGEMA 1036
DB 1020 DS-SHLEGAREEDGEVA 1035

RESULT 9
AAR89314
ID AAR89314 standard; Protein; 1060 AA.
XX
AC AAR89314;
XX
DT 01-JUN-1999 (first entry)
XX
DE SIIVmac239 genome pol protein sequence.
XX
KW Antigenic composition; primate; lentivirus; nef gene; vaccine;
KW infection; AIDS; SIIVmac239; pol protein.
XX
OS Simian immunodeficiency virus.
XX
PN US5851813-A.
XX
PD 22-DEC-1998.
XX
PF 27-JAN-1994; 94US-0188583.
XX
PR 27-JAN-1994; 94US-0188583.
PR 12-JUL-1990; 90US-0551945.
PR 09-JUL-1991; 91US-0727494.
XX

```


CC	(Updated on 25-MAR-2003 to correct PF field.)									
CC	(Updated on 25-MAR-2003 to correct PR field.)									
CC	(Updated on 25-MAR-2003 to correct PA field.)									
XX										
SQ	Sequence	1056	AA;							
<hr/>										
	Query Match	84.8%;	Score 4624;	DB 9;	Length 1056;					
	Best Local Similarity	81.9%;	Pred. No. 0;							
	Matches 849;	Conservative 91;	Mismatches 78;	Indels 18;	Gaps 1;					
QY	1	TGFRFTPLGKAPQLPRGPSSAGADTNTSPGSSSGSTGEIYAAREKTERAERTI	OG	60						
DB	39	TGFFRFPWPLGKEAPQFPHGSSASGADANCSPRTSCGSAKELHALGAAERKQRE	ALOG	98						
QY	61	SDRGLTAPRAGGDTIQGATNRGLAAPQFSILWRPVTYATIEGQPEVLLD	TGADDS	120						
DB	99	GD-----RGFAAPQFSILWRPVTYATIEGQPEVLLD	TGADDS	140						
QY	121	IELGNNYSKPIVGGIGGFINTREYKNEVEVLNKKVRATIMTGDTPINIFGRN	ILTALGM	180						
DB	141	IELGPHYTPKIVGGIGGFINTREYKNEVEVLNKKVRATIMTGDTPINIFGRN	LLTALGM	200						
QY	181	SLNLPVAKVEPIKIMLPGKDGPKLQWPLTKEKIEALKEICEKMEKQLEAP	TNPY	240						
DB	201	SLNLPVAKVEPIKIMLPGKDGPKLQWPLTKEKIEALKEICEKMEKQLEAP	TNPY	260						
QY	241	NTPTFAIKKKDKNKKRWLIDFRELNKVTDFTFQIOLGIPHPAGLAKKRIT	VLVDGDAYF	300						
DB	261	NTPTFAIKKKDKNKKRWLIDFRELNRVTQDTEVOIGIPHPAGLAKKRIT	VLVDGDAYF	320						
QY	301	SIPLHEDFRPYTAFTLPSVNNAPGKRYIKVLPQGWKSPALFOHTMQVLEP	FRKANP	360						
DB	321	SIPLDEFRQYTAFTLPSVNNAPGKRYIKVLPQGWKSPALFOHTMQVLEP	FRKANP	380						
QY	361	DVIIQYMDILIASRTDLEHVRVVLQKELLNGLFGSTPEKFOKDPYHWMG	YELWP	420						
DB	381	DVTLVQYMDILIASRTDLEHVRVVLQKELLNGLFGSTPEKFOKDPYHWMG	YELWP	440						
QY	421	TKWKLOKIQLOPEIKTVDNDIOLGVLANWAOLYPIGKTKHLIRKMLT	TEVQWT	480						
DB	441	TKWKLOKIQLOPEIKTVDNDIOLGVLANWAOLYPIGKTKHLIRKMLT	TEVQWT	500						
QY	481	ELAEAELENRIILSQEQEGHYQEKELEATVQKDQENQWYKIHQEEKILK	VGKYAKV	540						
DB	501	EMAEAEYENKIILSQEQEGHYQEKELEATVQKDQENQWYKIHQEEKILK	VGKYAKV	560						
QY	541	KNTHNGIRLLAQVQVQKIGKEALVINGRIPKPHLPVEREIVQWQWNYQV	TWIPDWFV	600						
DB	561	KNTHNGIRLLAQVQVQKIGKEALVINGRIPKPHLPVEREIVQWQWNYQV	TWIPDWFV	620						
QY	601	STPPLVRLAFNLVDPGPAETFYTDGSCNRQSKGAGYVTDGKDKVKKLE	QTTNOQA	660						
DB	621	STPPLVRLAFNLVDPGPAETFYTDGSCNRQSKGAGYVTDGKDKVKKLE	QTTNOQA	680						
QY	661	ELEAFAMALTDSGPKNIIVDSQVVMGIIISASOPTESKIVNQIIEEMIK	KEAIVYAWVP	720						
DB	681	ELEAFAMALTDSGPKNIIVDSQVVMGIIISASOPTESKIVNQIIEEMIK	KEAIVYAWVP	740						
QY	721	AHKGIGGNEVDHLVSGQIRQVLFLEKIEPAQEHEKYSNVKELSHKFC	GINLVARQIV	780						
DB	741	AHKGIGGNEVDHLVSGQIRQVLFLEKIEPAQEHEKYSNVKELSHKFC	GINLVARQIV	800						
QY	781	NSCAOQOQGEALHGVNAELGTWQMDCTHLEKILIVAVHVASGFIEA	EVIPQESGROT	840						
DB	801	DTCDKCHQGEALHGVNAELGTWQMDCTHLEKILIVAVHVASGFIEA	EVIPQESGROT	860						
QY	841	ALFLLKLASRWPTLHLDNGANFTSQEYKVMVAWWTGIEQSFQVYNP	QSQGVVEAMNHH	900						
DB	861	ALFLLKLASRWPTLHLDNGANFTSQEYKVMVAWWTGIEQSFQVYNP	QSQGVVEAMNHH	920						
QY	901	LKNQISRIEQAQNTIETIVLMAHCNMFRRGGIDMTSPERLINMITTE	QEQIQLQAKN	960						
DB	921	LKNQISRIEQAQNTIETIVLMAHCNMFRRGGIDMTSPERLINMITTE	QEQIQLQAKN	980						

QY	961	SKLKDFRVYFREGRODLWKGPBELLWKGGAVLVKVGTDIKIIPRRKAKIIRDYGGROQM	1020	
DB	981	SKFKNFRVYFREGRODLWKGPBELLWKGGAVLVKVGTDIKIVVPRRKAKIIDYGGGKEM	1040	
QY	1021	DSGSHLEGAREDEGEMA	1036	
DB	1041	DSSSHMEDTGEAREVA	1056	
RESULT 11				
AA	R22366	standard; Protein; 1060 AA.		
XX	AA	R22366;		
DT	06-JUL-1992	(first entry)		
DE	STVmac239	pol gene product.		
KW	Macaque; monkey; mac239;	Reverse transcriptase; retrovirus;		
KW	polymerase.			
XX	Simian immunodeficiency virus.			
OS	WO9200987-A.			
PN	23-JAN-1992.			
PD	10-JUL-1991;	91WO-US04884.		
PF	12-JUL-1990;	90US-0551945.		
PR	(HARD) HARVARD COLLEGE.			
PA	Desrosiers RC;			
XX	WPI; 1992-056816/07.			
PI	N-PSDB; AAQ22487, AAQ24802.			
XX	Primate lentivirus vaccine protecting against AIDS - and primate			
PT	lentiviruses and their DNA clones contg. null mutations, useful for			
PT	producing vaccine			
XX	Disclosure; Fig 1; 51pp; English.			
PS	Cell-free serum samples from a macaque monkey exhibiting symptoms			
XX	characteristic of SIV infection were co-cultivated with Hur-78			
CC	cells. Infectious SIVmac239 virus was identified in the cell			
CC	supernatant. Total cell DNA was prepared from SIVmac239-infected			
CC	cells and digested with EcoRI. An EMBL-4 library was constructed			
CC	from 10-20kb EcoRI fragments (EcoRI is a non-cutter of SIVmac239).			
CC	The library was screened with pK2 BamA as probe and a full-length			
CC	molecular clone was isolated and sequenced. This protein is the			
CC	pol gene product deduced from the proviral ORF.			
XX	Sequence	1060 AA;		
SQ	Query Match	84.7%; Score 4616; DB 13; Length 1060;		
	Best Local Similarity	81.8%; Pred. No. 0;		
	Matches	851; Conservative	88; Mismatches	79; Indels
			22; Gaps	2;
QY	1	TGFRFTGTLGKEAPQLPRGPSSAGADTNTSPGSSSGSTGEIY----	AAREKTERAERE	56
DB	39	TGFFRFPWSMGKEAPQFPHGSSASGADANCSPRSCGSAKELHAVGQAAERKQRE	98	
QY	57	TTGSDRGLTAPRAGGDTIQGATNRGLAAPQFSILMKRPVVTYATIEGQPEV	LLDTGADDS	116
DB	99	ALQGGD-----RGFAAPQFSILMKRPVVTYATIEGQPEV	LLDTGADDS	140
QY	117	IVAGIELGNNYSKPIVGGIGGFINTREYKNEVEVLNKKVRATIMTGDTPIN	IFGRNILT	176
DB	141	IVTGIELGPHYTPKIVGGIGGFINTREYKNEVEVLNKKVRATIMTGDTPIN	IFGRNILT	200

XX	AIDS; immunogen; antigen; vaccine; diagnostic.
KW	Simian immunodeficiency virus.
XX	WO80805440-A.
OS	28-JUL-1988.
XX	15-JAN-1988; 88WO-EP000025.
PN	16-JAN-1987; 87US-0003764.
PD	11-FEB-1987; 87FR-0001739.
XX	15-APR-1987; 87FR-0005398.
XX	(INSP) INST PASTEUR.
PA	(ALIZ/) ALIZON M.
PA	(CNRS) CENT NAT RECH SCI.
XX	Alizon M, Montagnier L, Guetard D, Clavfl F, Sonigo P, Guyader M;
PI	Tiollais P, Chakrabarti L, Desrosiers R;
XX	WPI; 1988-220290/31.
DR	New peptide(s) with immunological properties of HIV-2 envelope protein -
XX	having the structure of simian immune deficiency virus proteins,
PT	useful in diagnosis and of vaccine components
PT	Disclosure; Fig 5; 86pp; French.
XX	The SQ of pol protein of HIV-2 ROD (AAP80810) was compared with the SQ
CC	of pol protein of SIV MAC (AAP80809) in Fig 5. New peptides which have
CC	immunological properties in common with those of the peptide skeleton of
CC	the envelope protein of HIV-2 and also have a peptide structure in
CC	common with that of SIV-1 glycoprotein are claimed. Antigenic and
CC	immunogenic conjugates contg. the peptides and a kit to detect HIV-2 in
CC	biological fluids are new. The peptides are useful for in vitro diagnosis
CC	of HIV-2 infection and some of them can be used as components of
CC	immunogens and vaccines against HIV. Antibodies raised against them can
CC	be used for treatment of AIDS.
CC	(Updated on 25-MAR-2003 to correct PF field.)
CC	(Updated on 25-MAR-2003 to correct PR field.)
CC	(Updated on 25-MAR-2003 to correct PA field.)
XX	
SQ	Sequence 1056 AA;
	Query Match 84.3%; Score 4598; DB 9; Length 1056;
	Best Local Similarity 81.7%; Pred. No. 0;
	Matches 846; Conservative 89; Mismatches 83; Indels 18; Gaps 1;
Qy	1 TGRFFRTPLGCKEAPQLPRGPSSAGADTNTPSGSSSGTGEIYAAREKTERAERETIQG 60
Dd	39 TGGFEEFWPLGKEAPOPHGSSASGADANCSPRTSCGSAKELHALGOAAEKQREALQG 98
Qy	61 SDRLGTAPRAGDGTIOGATNRGLAAPQFSLMKRPVWTAYIEGQPVEVLDTGADDSIVAG 120
Dd	99 GD-----RGFAAPQPSLWRPVPVTAHIEGQPVEVLDTGADDSIVTG 140
Qy	121 IEIGNNYSPIKVGIGGFINTKEYKNVEIEVLNKVKVRATITMGDPINIFGNILTAGLM 180
Dd	141 IELPHYTPKTIVGGIGGFINTKEYKNVEIEVLGKRIRKGTITMGDPINIFGNILTAGLM 200
Qy	181 SLNLPAKVPEPIKIMLPKGDKPKLRQWPLTKKEIKTEALKEICEKMEKEGLEEAPTNPY 240
Dd	201 SLNLPIAKVEPVKSPLPGKDGPKLKQWPLSKKEITVALREICEKMEKDGQLEEAAPTNPY 260
Qy	241 NTPTFALKKDDKNRWMLIDREFLNKVTQDTEIOLGIPHAGLAKKRITVLDVDGDYAF 300
Dd	261 NTPTFALKKDDKNRWMLIDREFLNRVTDTEFVOLGIPHPAGLAKKRITVLDIGDAYF 320
Qy	301 SIPLHEDFRPYTFTPLFSVNNAEPGKRIYYKVLPOGWKGSIPAIFQHTMRQVLEPFRRANK 360
Dd	321 SIPLDEEFROYATTPLFSVNNAEPGKRIYYKVLPOGWKGSIPAIFYOYTHRWLVLEPFRRANK 380

AC AAP80809;

XX
DT 25-MAR-2003 (updated)

XX
ACCT NOV 27 1968

DE Sequence of pol protein of SIV isolate STLV-III Mac 142-83 in
DE clone SIV-1.

QY 361 DVIIIOYMDILLASDRDLEHVRVYLQKELLNGLFSTPDEKFOKDPYHMMGYELWP 420
 DB 381 DVTLVQYMDILLASDRDLEHVRVYLQKELLNGLFSTPDEKFOKDPYHMMGYELWP 440
 QY 421 TKWKLQIQLPQKEITWVNDIOKLGVLNWAAQYFGIKTKHLCLRLIRGKMTLITEVQWT 480
 DB 441 IKWKLQIQLPQKEITWVNDIOKLGVLNWAAQYFGIKTKHLCLRLIRGKMTLITEVQWT 500
 QY 481 ELAEAELEENRIILSFOEGHYHQQEKELEATVQKDOENQWTKIHQEEKILKVGKAKV 540
 DB 501 EMAEAEYENKIILSFOEGHYHQQEKELEATVQKDOENQWTKIHQEEKILKVGKAKV 560
 QY 541 KNTHTNGIRLLAOWKIGKEALVINGRIPKPHLPVEREITWQWVNDYQWVNDYQWVNDY 600
 DB 561 KNTHTNGIRLLAOWKIGKEALVINGRIPKPHLPVEREITWQWVNDYQWVNDYQWVNDY 620
 QY 601 STPLVRLVNLVDPFGAETFTYDGSNRSQKQKAGVYTDGKDKVKKLEQTTNQQA 660
 DB 621 STPLVRLVNLVDPFGAETFTYDGSNRSQKQKAGVYTDGKDKVKKLEQTTNQQA 680
 QY 661 ELAEAFAMALTDSPKYNIIIVDSQYVNGISASQPTSESKIVNQLIIEEMIKKEAIYVAVP 720
 DB 681 ELAEAFAMALTDSPKYNIIIVDSQYVNGISASQPTSESKIVNQLIIEEMIKKEAIYVAVP 740
 QY 721 AHKGIGGNORVHLVSGIRQVLFLEKIEPAQBEHEKYHSNKKELSHKFGIPNLVARQIV 780
 DB 741 AHKGIGGNORVHLVSGIRQVLFLEKIEPAQBEHEKYHSNKKELSHKFGIPNLVARQIV 800
 QY 781 NSCAQCOQKEATHGQVNAELGWQMDCTHLEKIIIVAVHVASGFIEAEVQESGROT 840
 DB 801 DTCDKCHQGEATHGQVNSDLGTWQMDCTHLEKIIIVAVHVASGFIEAEVQESGROT 860
 QY 841 ALFLLKLASRPITLHTONGANFTSOEVKMWAWWJGIESQGVYVNPQSGVVEAMNHH 900
 DB 861 ALFLLKLASRPITLHTONGANFTSOEVKMWAWWJGIESQGVYVNPQSGVVEAMNHH 920
 QY 901 LKNOISRIEQAANTIEIVLMAHCFNRRGIGDMTPSERLINMITTEQETQFLOAKN 960
 DB 921 LKNOISRIEQAANTIEIVLMAHCFNRRGIGDMTPSERLINMITTEQETQFLOAKN 980
 QY 961 SKLKDPRVYFREGDQWLGPGELLWKGAGVILVKVGTDIKIIPRKAIIIRDYGGQEM 1020
 DB 981 SKLKDPRVYFREGDQWLGPGELLWKGAGVILVKVGTDIKIIPRKAIIIRDYGGQEM 1040
 QY 1021 DSGSHLEGAREDEGEMA 1036
 DB 1041 DSSSHMEDIGAREVA 1056
 RESULT 13
 AAB12993
 ID AAB12993 standard; Protein; 1055 AA.
 XX AAB12993;
 AC AAB12993;
 XX AAB12993;
 DT 30-NOV-2000 (first entry)
 XX HIV-2 reverse transcriptase protein sequence SEQ ID #39.
 DE HIV-2 reverse transcriptase protein sequence SEQ ID #39.
 XX Reverse transcriptase; RT; polymerase chain reaction; PCR; retrovirus;
 KW stability; solubility.
 XX Human immunodeficiency virus type-2.
 OS Human immunodeficiency virus type-2.
 XX WO200042199-A1.
 XX 20-JUL-2000.
 XX 14-JAN-2000; 2000WO-US00896.
 PF 15-JAN-1999; 99US-0116099.
 PR 15-JAN-1999; 99US-0116099.
 XX

(MOLE-) MOLECULAR BIOLOGY RESOURCES.
 Swaminathan N;
 WPI; 2000-482830/42.
 Novel genes encoding reverse transcriptase polypeptides modified by altering or adding the integrase domains by truncation internally and/or at the C-termini, useful in cDNA synthesis and amplification procedures -
 Claim 1; Page 142-146; 189pp; English.
 This invention relates to isolated polynucleotide sequences encoding a polypeptide with RNA dependent DNA polymerase (reverse transcriptase RT) activity. RTs are found in a variety of retroviruses and their defining activity is the ability to synthesise a cDNA strand using an RNA template. The invention includes sequences AAA87808-A87822 and cDNA AAA87840-A87842 which represent nucleotide sequences encoding RT polypeptides. Included in the invention are reverse transcriptase protein sequences AAB12989-B12995. Sequences AAA87829-A87839 and AAA87843-A87897 represent oligonucleotides used in the identification and synthesis of the RT nucleotide sequences of the invention. Many of the RT nucleotide sequences encode modified RT proteins, which exhibit improved stability and/or improved solubility, relative to naturally occurring reverse transcriptases. The modified RT nucleotide and protein sequence are used in improved polymerase chain reaction (PCR) methods. The nucleotide sequences can be used in sequencing methods.
 The present sequence represents a reverse transcriptase protein of the invention.
 Query Match 82.9%; Score 4519.5; DB 21; Length 1055;
 Best Local Similarity 79.7%; Pred. No. 0;
 Matches 839; Conservative 93; Mismatches 86; Indels 35; Gaps 3;
 Sequence 1055 AA;
 1 TGRFFRTPLGKEAPQLPR-----GPS-----SAGADTNTPSGSSSGTGEI 43
 18 TGGFFVRPLGKEASQFPFGDPSAICAPDEPSIRHDTSGDSDICTPCRSSRGDAKEL 77
 44 YAAREKTERAERETQGSDRGLTAPRAGDITQATNRLAAPQSLKRPVVTAYIBGO 103
 78 HATREAEQERETLQGGD-----RGFAAPQSLKRPVVTAYIBGO 119
 104 PVEVLDTGADDSIVAGIELGNYSKPIVGGIGGINTKEYKNVIEVLNKKVRATIMTG 163
 120 SVEVLDTGADDSIVAGIELGNYSKPIVGGIGGINTKEYKNVIEVLNKKVRATIMTG 179
 164 DTPINIFGRNIIALGMSLNLPAKVEPTIKMLPKGDKGPKLRWPLTKKEIKELKEICE 223
 180 DTPINIFGRNIIALGMSLNLPAKVEPTIKMLPKGDKGPKLRWPLTKKEIKELKEICE 239
 224 KMEKEGLEEAPPTNPYPTFAIKKKDKNKKRMILDFRELNVQDTFETQLGPHFAG 283
 240 KMEKEGLEEAPPTNPYPTFAIKKKDKNKKRMILDFRELNVQDTFETQLGPHFAG 299
 284 LAKKRITVLDVGDAYFSIPLHEDFRPYTAFTLPVNNAPGKRIYKVLPGWKGSPAI 343
 300 LASKRIITVLDVGDAYFSVPLDPFRQYTAFTLPVNNAPGKRIYKVLPGWKGSPAI 359
 344 FOHTMRQVLEPFRKANKDVIIOYMDILLASDRDLEHVRVYLQKELLNGLFSTPDE 403
 360 FOYTMKVLDPFRKANKDVIIOYMDILLASDRDLEHVRVYLQKELLNGLFSTPDE 419
 404 FKQDPPPHWGYELWPKKWLQKIQLPKELTWVNDIOKLGVLNWAAQYFGIKTKHL 463
 420 FKQDPPPHWGYELWPKKWLQKIQLPKELTWVNDIOKLGVLNWAAQYFGIKTKHL 479
 464 CRLIRGKMTLITEVQWTELAELEENRIILSFOEGHYHQQEKELEATVQKDOENQWTK 523
 480 CRLIRGKMTLITEVQWTELAELEENRIILSFOEGHYHQQEKELEATVQKDOENQWTK 539

QY 524 KIHQEKILKVGKYAKVKNTHNGIRLLAQVVKIKGKALYVWGRIPKPHLPVERIEWEQ 583
 Db 540 KIHQDKILKVGKYAKVKNTHNGIRLLAHVVKIKGKALYVWGRIPKPHLPVERIEWEQ 599
 QY 584 WNDYVQVWTVLPDWDVSTPLVRLAFNLVGDPIPGAETFFYDGSNCRQSGKAGYVTD 643
 Db 600 WMTDYVQVWTVLPDWDVSTPLVRLAYNLVDPLEGVETTYDGSNCRQSGKAGYVTD 659
 QY 644 RGKDKVKELETTNQAELEAFAMALTDGSPKVNIVDSQVVMGISAQPTSESKIVNQ 703
 Db 660 RGKDKVKELETTNQAELEAFALQDSGPQVNVIVDSQVVMGIVAQAQPTSESPVRE 719
 QY 704 IIEEMIKKEIYVAVVPAHKIGGQEVVDHLVSGIROVLFLEKTEPAQEEHEKYHSNVK 763
 Db 720 IIEEMIKKEIYVGVVPAHKGLGGNQEVVDHLVSGIROILFLEKTEPAQEEHEKYHNVK 779
 QY 764 ELSHKEFGPNLVARQIVNSCAQCKGEAIGHQVNAELGTWQMDCTHLEGKIIIVAVHVA 823
 Db 780 ELVHKFGIPQLVARQIVNSCDKQCKGEAIGHQVNSSELGTWQMDCTHLEGKVIIVAVHVA 839
 QY 824 SGFTEAEVPOESGRQTALFLKLASRPITHLHTDNGANFTSQEVKVMVWVWIGIEQSGF 883
 Db 840 SGFTEAEVPOETGRTALFLKLASRPITHLHTDNGANFTSQDVKMAWVWIGIEQTFG 899
 QY 884 VYPNPQSGVVEAMNHHKNOISRIEQANTETIVLMAIHCNMFRRGGIGDMPSPRL 943
 Db 900 VYPNPQSGVVEAMNHHKNOIDRIDQAVSTETVLMATHCMNFKRRGGIGDMPSPERI 959
 QY 944 INMLTTEQEIOLFQAKNSKLDKFRVYFREGDQLWKGPGELLWKGEGAVLVKVGTDIKII 1003
 Db 960 VNMILTTEQEIOLFQAKNLKFNFRVYFREGDQLWKGPGDLLWKGEGAVLVKVGTEIKVI 1019
 QY 1004 PRRAKIIIRYDGGROEMSGSHLEGAREDEGMA 1036
 Db 1020 PRRAKIIIRYGGKELDCSADVEDTQAREVA 1052
 RESULT 14
 AAB01769
 ID AAB01769 standard; Protein; 1018 AA.
 XX AAB01769;
 AC AAB01769;
 XX 03-JAN-2001 (first entry)
 DT Simian immunodeficiency virus SIVrcm Pol protein #1.
 DE Red capped monkey; simian immunodeficiency virus; SIV; SIVrcm;
 XX Red vaccination; Pol protein.
 KW Simian immunodeficiency virus.
 OS
 XX Key Location/Qualifiers
 FT Misc-difference 569
 FT /note= "encoded by GACTAA"
 XX WO200034529-A1.
 PN 15-JUN-2000.
 PD
 XX 03-DEC-1999; 99WO-US28638.
 PF 07-DEC-1998; 98US-0206551.
 PR (UABR-) UAB RES FOUND.
 PA Hahn BH, Shaw GM, Gao F, Marx PA, Smith SM, Georges-Courbot MC;
 XX Lu CY;
 PI WPI: 2000-423454/36.
 DR N-PSDB; AAA51008, AAB01775, AAB01776.
 XX New polypeptide, useful for producing antibodies and for diagnosis of

PT simian viral infection in humans comprises complete genome of new
 PT simian immunodeficiency virus isolate -
 XX Disclosure; Page 152-156; 173pp; English.
 PS
 XX The present sequence is the Pol protein from the simian immunodeficiency
 CC virus found in the red capped mangabey and designated SIVrcm. This virus
 CC is related to the HIV viruses that cause AIDS in humans. Knowing the
 CC sequence of the Pol protein and its gene is useful as it enables
 CC screening for the virus to take place in humans and animals (there is a
 CC possibility that this virus undergoes cross-species transmission), allows
 CC vaccines to be produced and aids research into the origin of these
 CC viruses. This latter is important if an AIDS vaccine is to be found.
 XX
 SQ Sequence 1018 AA;
 Query Match 61.0%; Score 3324; DB 21; Length 1018;
 Best Local Similarity 59.5%; Pred. No. 4.le-258;
 Matches 627; Conservative 156; Mismatches 209; Indels 62; Gaps 8;
 QY 4 FRTGTPLCK-EAPOL-----PRGPSSAGADTNPSPSGSGST 40
 Db 1 FRETPLSGGEAEELSFDFLSDCRDGEQLRPCRDAAKELSEERGKTETE----- 53
 QY 41 GEIYAAREKTERAERETIQGSDRGLTAPRAGGDTIQGATNRGLAAPQFSLMKRPVYTAI 100
 Db 54 ----AGRQEER-----GSISGSLNLPQFALMKRPPTIAQI 85
 QY 101 EGQPVVELLDTCADDISIVAGIELGNYSKPIVGGIGGINTKEYKNVIEVLNKKVRATY 160
 Db 86 EGQKVEVLDTGADDTVIEGIELGNDWTPKIIGGIGGYINVKYKNCIEIAGKRTAHAV 145
 QY 161 MTGDTPIINFGRNILTALGSLNLPVAKVEPTIKMLKPKGDKGPKLROPLTKETALKE 220
 Db 146 LVGPTYPNIIIRNVLKLGATLNFPSIQIETIKVELKSGQDGPVRYKPLSKETALTE 205
 QY 221 ICEKMEKEGQLEAPPTNPYNTPTPAIKKKDKNKRMLIDFRELKNVQDTEIOLGIPH 280
 Db 206 ICNAMEKEGKISKIGPENAYPTPIFCIKKOSTKWKRLVDPRFELNKRQDFFEVOLGIPH 265
 QY 281 PAGLAKKRITVLDVGDAFYSIPLHEDFRPYTAFTLPSVNAEPKRYIYKVLPGWGKS 340
 Db 266 PGLKQCEIRITVLDIGDAFYFCLLYEPFRKYTAFTIPAVNNQGGPVRYQYVLPQWGKS 325
 QY 341 PAIFQHTMROVLEPERKANKDVIIIOYMDLILIASDRDLDLHEDRVVLQKELLNGLGFS 400
 Db 326 PAIFQHSANKILQPPREENPDVYIYQYMDLDFVSDRDKLEHDKMICKOLRDLHLLFWGET 385
 QY 401 PDERFKDPPYHMGYELWPTKWKIQIKQLPQKEITWTVNDIOKLGVNLNAAQLYPGKRT 460
 Db 386 PDKKFDKPPYLLWGYELHLPKSWTVQEIKLPEKEEMTVNDIOKLGVNLNAAQLYSGLRT 445
 QY 461 KHLCLRLIRGKMTLPEEVOWTELAELAEENRIILSOBOEGHYQYEEKELEATVQKQDNQ 520
 Db 446 KELCLIRGAKALDEKVENTEAEIYEENKMLKELHGVYIYDEKPLVANIQLLEGQ 505
 QY 521 WYKITHQEE-KILKVGKYAKVKNTHNGIRLLAQVVKIKGKALYVWGRIPKPHLPVER 579
 Db 506 WSYQIEGSGPLKTGKYAKOKTAHTNEIRMLAGLVQKIAKEAIVWGRIPKPHLPVER 565
 QY 580 IWEQWMDNYWQVWTVLPDWDVSTPLVRLAFNLVGDPIPGAETFFYDGSNCRQSGKAG 639
 Db 566 WVD-WRSQYVQVWTVLPDWEFVSTPLIRGLYNLVKDPPIGEEVYVVDGAAARNNSKIGK 624
 QY 640 YVTORGDKVKKLEQTTNQAELEAFAMALTDGSPKVNIVDSQVVMGISAQPTSESK 699
 Db 625 YVTNRGKVKLELETTNQAELEAVLLALAKDSGPKVNIVDSQVYVILEAQPTSDSG 684
 QY 700 IVNQIIEEMIKKEIYVAVVPAHKIGGQEVVDHLVSGIROVLFLEKTEPAQEEHEKYH 759
 Db 685 LVTEIINQIMIGKEAVYLSVFAHKIGGNEEVVDKLVSKGIRQVLDLGGIEKAEQEEHEKYH 744
 QY 760 SNVKELSHKFGIPNLVARQIVNSCAQCKGEAIGHQVNAELGTWQMDCTHLEGKIIIVA 819

XX	Alizon M, Montagnier L, Sonigo P, Wain-Hobson S;				
PI	WPI: 1998-582548/49.				
DR	N-PSDB; AAV63467.				
XX	Human immunodeficiency virus isolate LAV(MAL) - and method for				
PT	detecting anti-HIV antibodies				
XX					
PS	Disclosure; Fig 7B-E; 47pp; English.				
XX					
CC	This is the amino acid sequence of the pol protein of				
CC	lymphadenopathy associated virus LAV.MAL (CNCM I-641), a new virus				
CC	isolate from Zaire that is responsible for diseases clinically				
CC	related to AIDS. The sequence was deduced from an open reading				
CC	frame (ORF) of the LAV.MAL genome (see AAV63467). 7 ORFs (see				
CC	AAW72992-98) were identified. Specific peptides of the envelope				
CC	glycoprotein can be used as antigens in a claimed method for the in				
CC	vitro detection of an antibody directed against LAV. The method is				
CC	useful for the diagnosis of AIDS or pre-AIDS, or to detect				
CC	antibodies in patients, asymptomatic carriers and in blood-related				
CC	products. LAV viral antigens are also useful in vaccines.				
CC	(Updated on 25-MAR-2003 to correct PR field.)				
XX					
SQ	Sequence 1002 AA;				
Query Match 57.9%; Score 3157; DB 19; Length 1002;					
Best Local Similarity 57.3%; Pred. No. 1.1e-244;					
Matches 576; Conservative 177; Mismatches 236; Indels 16; Gaps 5;					
QY	32	PSGSSSGSTGEIYAAREKTERARETIOGSDRLGAPRAGGDTIOGATNRGLAA---PQF 88			
DB	9	POCKAREFPSEQTRANSPTSRELR--VWGDKTLSE-----TGAERQGIVSFFPQI 58			
QY	89	SLWKRPPVTAIEGQPEVLLDGCADDSIVAGTELGNNSPKIVGGIGGIFINTKEYKNVE 148			
DB	59	TLQWRPVTVRVGGQLKEALLDGTADDTVLEINLPGKWKPKMIGGIGGIFKVRQDQIL 118			
QY	149	IEVLNKKVRAITWGTPINIFGRNLTALGMSLNLPVAKVEPIKIMLPGKDGKPLRW 208			
DB	119	IEICGKAIGTILVGTPTNIGRNMLTQIGCTLPNPIETVPVKLPKPGMDGPRVKOW 178			
QY	209	PLTKETIEALKEICERMEKEGLEEAPPNTPTFAIKKKDKNKRMLIDPRELNKVT 268			
DB	179	PLTEEKIKALTEICKMEKEGKILKIGPENPYNTPVFAIKKDKTKRKLNVNRELNKT 238			
QY	269	ODTFETQLGHPHAGLAKRRITVLDVGDAYSIPLHEDFRPYTATPLPSVNNAEPGRY 328			
DB	239	QDFEVLQGLGHPHAGLAKKSVTLVDVGDAYSFVLDEDFRKYTFTIPSINNETPGIRY 298			
QY	329	IYKVLPGQWKGSPAIFQHTMROVLEPFRKANKDVIIIOYMDDLIASDRDLEHDRVVLQ 388			
DB	299	QYNVLPQWKGSPAIFQSSMTKILEPFRKNPBIYQYMDDDIYVGSDDLEIGHRTKIEE 358			
QY	389	LKELLNGLGFSTPDEKFDKPPYHMMGYELWPTKWKLOKIQLPQKEIWTNDIQKLVGL 448			
DB	359	LREHLKMGFTTDDKKHQKEPPLMWGYELHPDKVTVQIPLQPDKESVTNDIQKLVGL 418			
QY	449	NWAAQIYPGKTKHLRLRGKMTLTFEVOVWTELAEAELEENRIILUSQEGEHYQEEKE 508			
DB	419	NWASQIYPGKTKVQLCKLRGAKALTDIVPLTAEAELEAENREILKEPVHGVYDPSKD 478			
QY	509	LEATVOKDQENOWTKIHQEE-KILKVGKVAKNTHNTGIRLLAQVQVQKIGKEALVWG 567			
DB	479	LIAEIQKQGGQWYQIYQEQYKNTKGTARIKSAHTNDVKOLTEAVOKIAQESIVWG 538			
QY	568	RIPKFLPVEREIQEOWBNYQVWIPDWDVSTPPLVLAFLNVLGDPDIPGAETFYTDG 627			
DB	539	KTPKFLPLIQKETWEAWWTEYQATWIPWEFVNTPPLVKLVQLETEPIVGAETFYVDG 598			
QY	628	SCNRQSEKAGVITDRGDKVKKLEBQTTNQAELEAFAMALTDGSPKWNIIIVDSQVYG 687			
DB	599	AANRETKKAGYVITDRGRQVSVLTETTNQKTELQAIHLALQDSGSEVNIYDTSQYALG 658			

QY	688	ISASQPTSESKIVNOIIEEMIKKEAIYVAVWPAHKGIGGNQEVHVLVSQIGROVLFLEK 747			
DB	659	IQAQDPKSESEIVNOIIEQLIQKDVILSWWPAHKGIGGNQEVKLVSSGIRKVLFLDG 718			
QY	748	IEPAQEEHEKYHSNVKELSHKFGIPNLVARQIVNSCAQQOQKGEAIGHQVNAELGTWMD 807			
DB	719	IDKAOEHEKYHSNWRAMASDFNLPIVAKEIVASCDKQLKGEAMHGQVDCSPGIWLD 778			
QY	808	CTHLEKIIIVVHVHVASGFIEAEVPOESGRQTALFLKLKASRWPITHLHTONGANFTSQ 867			
DB	779	CTHLEKIIIVVHVHVASGYIEAEVPAETQOETAYEILKLGRWPKVYVHTONGSNFTSA 838			
QY	868	EYKMWAAWIGIEQSGFQVPYNSQGVVEAMNHLKXNOISRIREQANTIEFIVLMATHCMN 927			
DB	839	AVKAAACWANIYQEFQIPYNSQGVVESMKNELKIIQGVREQAELHKTAYQMAVFIIN 988			
QY	928	FKRRGGIGDMTPSERLINMITTEQEIQFLOAKNSKLKDFRVYFREGRDQLWKGPGELLWK 987			
DB	899	FKRKGIGGYSAGERIIDMIATDIQTKELQKQITKIQNFRVYVRDNRDPWKGPAKLLWK 958			
QY	988	GEGAVLVKVTGDIKIIPRKAKIIRDYGGROEMDSGSHLEGARED 1032			
DB	959	GEGAVVIQDNSDIKVVPRRKAKIIRDY--KOMAGDDCVAGGODE 1001			
RESULT 18					
AAW68474					
ID	AAW68474	standard; Protein; 1014 AA.			
XX	AAW68474;				
XX	AC				
DT	25-MAR-2003	(updated)			
DT	08-DEC-1998	(first entry)			
XX	HIV-1 strain YBF30	pol protein.			
DE	HIV-1 strain YBF30;	antibody; oligonucleotide; diagnosis; immunisation;			
KW	HIV-1 strain YBF30;	infection; typing; pol.			
XX	Human immunodeficiency virus type 1.				
OS	FR2756843-A1.				
PN	12-JUN-1998.				
XX					
PD	09-DEC-1996;	96PR-0015087.			
XX	09-DEC-1996;	96PR-0015087.			
PR	(ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.				
XX	(INRM) INSERM INST NAT SANTE & RECH MEDICALE.				
PA	(INSP) INST PASTEUR.				
XX					
PI	Mauclere P, Lousserat AI, Simon F, Saragosti S, Barre SF;				
XX					
XX	WPI: 1998-336114/30.				
DR	N-PSDB; AAV60754.				
XX	Non-M, non-O HIV-1 strain YBF30 - useful for diagnosis and				
PT	immunisation				
PS	Claim 3; Page 29-33; 85pp; French.				
XX					
CC	This sequence represents the pol protein from the non-M (major), non-O				
CC	(Outlier) HIV-1 strain YBF30 (CNCM I-1753), isolated from the Cameroon.				
CC	The HIV strain (see AAV60751 for complete genome), peptides, antibodies				
CC	and oligonucleotides derived from it (see AAV60752-V60798 and				
CC	AAW68473-W68482) are used for diagnosis of or immunisation against non-M,				
CC	non-O HIV-1 infections. The oligonucleotides, peptides and antibodies				
CC	can also be used for typing HIV strains.				
CC	(Updated on 25-MAR-2003 to correct PI field.)				
XX					

SQL	Sequence	1014 AA;
	Query Match	57.8%; Score 3150.5; DB 19; Length 1014;
	Best Local Similarity	56.0%; Pred. No. 3.8e-244;
	Matches	579; Conservative 179; Mismatches 253; Indels 23; Gaps 6;
QY	4	FFRTGPGKAPQLPRGPSSAGADNTSPSSSGSTGEIYAAREKTERAERETIQ-GSD 62
DB	1	FFR-----EELVSLQRETRKLPDPNNK--EHAHPATRELWVSGGEHTGEGDAGEPGED 53
QY	63	RGLTAPRAGGDTIQGATNRLAAPQFSLWKPVPVYATIEGQPVVELLDTGADDSIVAGIE 122
DB	54	RELSVPT-----FNFQITLWQRPVITVKIGKVEALLDTGADDTVIELQ 100
QY	123	LGNNYSPKIVGIGGIFNTKEYKNVEIEVLNKKVRATIMTGDTPINIFGRNLTALGMSL 182
DB	101	LEGRWPKMIGGIGGIFKVRQYDNTVDIQKAVGTIVLGPVPTVNIIGRNLLTQIGCTL 160
QY	183	NLPVAKVEPIKMLKPGDKGPKLQWPLTKIEIKALKEICEKMEKEGQLEEAPPTNPNYNT 242
DB	161	NFPISPIETVPVKLPKPGMDGPKVQWPLTTEKIEALREICTEMEKEGKISRIGPENPNYNT 220
QY	243	PTFAIKKKDKNKKWMLIDFRELNVKVTQDFTETIQLGHPHAGLAKKRRITVLVDGDAYESI 302
DB	221	PIFAIKKKDSTKWRKLVDFRELNKRTODFWEVLGHPHAGLKQKSVTVLVDGDAYFSC 280
QY	303	PLHEDFRPYTAFTLPSVNNNAEPGRKRYIYKVLPGQWKSPAFIQHTMROVLPEPFRKANKDV 362
DB	281	PLDKRKYTAFTIPSNNETPGIRYOYNVLPGQWKSPAFIQSTMTKILPEFREKHPFI 340
QY	363	IIIQYMDLIIASRDTLEHRRVVLQKELNGLGTFSTPDEKFKQDPPYHWMGLWLPTR 422
DB	341	IIYQYMDLYVGSDELAQHREAVEDRLHLLKGFPTPDKKHQKEPPFLMWGYELHDPK 400
QY	423	WKLQKIQLPKEIITWVNDIOKLVGLNAAQLYPGIKTKHLCLRLGKNTLITEVQWTEL 482
DB	401	WTVQPIKLPKDWMTVNDIOKLVGLNAAQLYPGIRVKQJCKLIRGARALTVEYVNFTEE 460
QY	483	AEAELENRIITLSEQEGHYHQEKELEATVQKQENQWTKYKHQE-EKILKVGKYAKVK 541
DB	461	AELEAENREILKEPLHGVYDDPKGELVAELQKQOGQWYQIYQELHKNLTKGYAKMR 520
QY	542	NHTNGIRLLAQVQKIGKEALVWIGRIPKFLPVEREINWQMDNWNQVWTPDWDVPS 601
DB	521	SAHTNDIKOLVEVYRKVATESIVTWGTPKFLPQKEVWDAWNTDHWATWIDEPWEFVN 580
QY	602	TPPLVRLAFLNVLGDPICGAETFTYDGSNRSKESKAGYVTRDCKVKKLEQTNQAAE 661
DB	581	TPPLVRLWYQLETFISGAETFTYDGAANRETKLKGAGFVTRGRQKVVSADTNNQAAE 640
QY	662	LEAFAMALTDGPKVNIIVDSQYVNGISASOPTSESKIVNQIIEEMIKKEAIYVAVWPA 721
DB	641	LQAILMALQESGRDNIIVTDSQYAMGIHSPDKSESELVSQIIEELIKKERVLSWVPA 700
QY	722	HKGIGGNOVDHLVSGIRQVLFLEKTEPAQEEHEKHVSNVKELSHKFGIPNLVARQIVN 781
DB	701	HKGIGGNEQVDKLVSSGIRKILFLDGIEKAQEDHRYHSHNWKAMASDFNLPIVAKETVA 760
QY	782	SACQOQGRAIHGOVNAELGTQMDCDTHLEGKIIIVAVHVASGFIAEVIPOESGRQTA 841
DB	761	SCDCKQLKGEMHQVNCSPGVQWLDCTHLEGKIIIVAVHVASGFIAEVIPOESGRQTA 820
QY	842	LFLKLKASRPITHLHDTNGANFTSOEVKVMYMWIGIEQSGVPYNPQSQGVVWAMNHL 901
DB	821	YFILKLAGRVPKVIHTDNGSNFTSATVKAACWANIKQEEGIPYNPQSQGAVESMKNEL 880
QY	902	KNQISRTREQANTTETVLMAHCMNFKRRGGIDMTPTSERLINMITTEQBIQLOAKNS 961
DB	881	KKIIGQIRDOAEHLKTAQVMAVFTHNFKRRGGIGGYTAGERIIDIATIDQTNLQTOIL 940
QY	962	KLKDFRYFEREGRDOLWKGPCCELLWKGEGAVLVKVGTDIKIIPRKAIIIDYGGROEMD 1021
DB	941	KVQNFYVYRDSRDIWKGPAKLLWKGEGAVVYIQDNGDIKVVPRRKAKIIRDY-GKQWAG 999

QY	1022	SGSHLEGAREDEGM 1035
DB	1000	DGVASGQDENQEM 1013
RESULT 19		
ID	AAB69287	standard; Protein; 1005 AA.
XX	AAB69287;	
AC	AAB69287;	
DT	20-APR-2001	(first entry)
XX	HIV-1 non-subtype B clone 962M751-3	pol protein.
DE	HIV-1; human immunodeficiency virus; non-subtype B; gag; pol; env;	
KW	vpu; vif; vpr; tat; rev; nef; vaccine.	
XX	Human immunodeficiency virus type 1.	
OS	WO2000026416-A1.	
PN	11-MAY-2000.	
PD	25-OCT-1999;	99WO-US24837.
XX	02-NOV-1998;	98US-0184418.
XX	(UABR-) UAB RES FOUND.	
XX	Hahn BH, Shaw GM, Gao F;	
PI	WPI; 2000-365651/31.	
DR	Novel genomic nucleic acids of non-subtype B human immunodeficiency	
XX	virus type 1 useful for detecting and treating AIDS comprises a	
PT	specific nucleotide sequence	
PS	Claim 41; Fig 15; 131pp; English.	
XX	The present invention provides the protein and coding sequences for a	
CC	number of human immunodeficiency virus (HIV) type 1 non-subtype B	
CC	isolates. The sequences shown include the near full-length coding	
CC	sequences from each isolate, and the env, pol, vif, vpr, vpu, gag, tat,	
CC	rev and nef proteins. These can be used to detect the presence of HIV-1	
CC	in a sample and to produce antibodies against non-subtype B HIV-1 virus.	
CC	These antibodies can be used in vaccines to prevent and treat HIV	
CC	infection.	
XX	Sequence	1005 AA;
SQL	Query Match	57.5%; Score 3134.5; DB 21; Length 1005;
	Best Local Similarity	55.9%; Pred. No. 7.3e-243;
	Matches	577; Conservative 184; Mismatches 230; Indels 41; Gaps 9;
QY	4	FFRTG---PLGKEAPQLPRGPSSAGADNTSPSSSGSTGEIYAAREKTERAERETIQ 60
DB	1	FFRENLAPEEG-EGALPSEQTRA-----NSPTSSNPTSREL-----QVRG 41
QY	61	SDRGLTAPRAGGDTIQGATNRLAAPQFSLWKPVPVYATIEGQPVVELLDTGADDSIVAG 120
DB	42	DN---PCPEAGAER-QGTFLN---CPQITLMORPLVSIKVGQIKAEALLDTGADDTVLEE 93
QY	121	IELGNNYSPKIVGIGGIFNTKEYKNVEIEVLNKKVRATIMTGDTPINIFGRNLTALGM 180
DB	94	INLPCKWPKRMIGGIGGIFKVRQYDQILIEICGKKAIGTVLGVPTVNIIGRNMLTQJGC 153
QY	181	SLNLPVAKVEPIKMLKPGDKPKLQWPLTKIEALKEICEKMEKEGQLEEAPPTNPY 240
DB	154	TLPNPISEIETVPVKLPKPGMDGPRVKQWPLTEKIKALTAICEEMEKEGKTKIGPENPY 213
QY	241	NTPTFAIKKKDKNKKWMLIDFRELNVKVTQDFTETIQLGHPHAGLAKKRRITVLVDGDAYF 300

```

Db 214 NTPVFAKKKSTKWRKLVDFRELNKRQDQFEVQLGIPHPAGLKKKSVTVLVDGDAVF 273
Qy 301 SIPLHEDFRPYTAFTLPSVNNAEPAKRYIKYKVLPGQWKGSPAFQHTMRQVLEPRKANK 360
Db 274 SVPLDGFRRKYTAFTIPISINNETPGIRYQYNVLPQGWKGSPSIFQSSMIKILEPFRTPQ 333
Qy 361 DVIIQYMDLIIASDRTDLEHDRVVLQKELNGLGFGSTDPKFKDPPYHWMGYELWP 420
Db 334 EIVIQYMDLVGSDLEIGHRAKIEELRHLRWGFTTPDKKHQKEPFFLWNGYELWP 393
Qy 421 TKWKIQIOLPOKEIWTNDIQKLVGLNWAQLYPGIKTKHLRCLIRKMTLIEEVQWT 480
Db 394 DKWTVQPIKLPEKESVTNDIQKLVGLNWXSIYAGIKVRQKLLRGAKALTDIVPLT 453
Qy 481 ELAAELEENRIISQOECHYQOEKELEATVOKDOENQWYKHIOEE-KILKVGKYAK 539
Db 454 EEAELEABSEIREILKEPVHGVYDFSPKDLIAEQKGHDQWYQVYQEPFKNLTKGYAK 513
Qy 540 VKNTHNGIRLLAQVQKIGKEALVIGWGRIPFHLPPVEREIVEQWMDNWTWTPDWD 599
Db 514 MTAHTNDVKQLTEAVOKTAMESIVLWGIKIPERLDIQKETWETWTDIWAQWIPWEF 573
Qy 600 VSTPPLVRLAFNLVGDPIPAETFTYDSCNRQSKGKAGYVTDGKDKVKKLEQTNQ 659
Db 574 VNTPLVLKWLQLEKEPIAGAEYTVYVDAANRETKIGKAGYVTDGROKIVTLTETNOK 633
Qy 660 ALEAPAMALTDSPKPVNIIVDSQYVMGISASQTESKTYNQIIEEMKKEAYVAV 719
Db 634 TELQAIQLALQSGSEVNTVDSQALGIIQAQPKSESELVNIQIEQLIKERVLSW 693
Qy 720 PAHKGIGGNOVDHLVSGIROVLFLEKTEPAQEEHEKYHNVKELSHKFGIPNLVAROI 779
Db 694 PAHKGIGGNEQVDKLVSSGIRKVLFDGIDKAAQEEHEKYHNSWRAMASEFNLPPIVAK 753
Qy 780 VNSCAQCKQGEAIGQVNAELGTWMDCTHLEGKIIIVAVHVASGFIEAEPQESGR 839
Db 754 VASCDKQLKGEAIGQVDCSPGIWQDCTHLEGKVLVAVHVASGYIEAETPAETGOE 813
Qy 840 TALFLLKASRPITLHTDNCANTSOEVKVAWVWIGIEQSGVPYNPQSGVVEAMNH 899
Db 814 TALFLLKAGRPVVKVHTDNGSNFTSAAVKACWAGIHQEGIPYNPQSGVVEAMNK 873
Qy 900 HLKNOISIREQANTIETVLMAHCMNPKRRGGIGDMTPSPRLNMIITEOEIOFLQAK 959
Db 874 ELKKIIGVQDQAEHLKTAQVAVFTHNPKRGIGGYSAGERIIDIIATDIQTRQLQK 933
Qy 960 NSKLDKFRVYFREGDQLWKGCELLWKGEGAVLVKVGTDKIIPRRKAKIIRDYG --- 1015
Db 934 IIKIQNFRYVYRDSRDPWKGPAKLWKGEGAVVIQDNSDIKVIIPRRKAKIIRDYGQMA 993
Qy 1016 -----GROEMD 1021
Db 994 GTDSVAGRODED 1005

RESULT 20
AAB69288
ID AAB69288 standard; Protein; 1003 AA.
XX
AC AAB69288;
XX
DT 20-APR-2001 (first entry)
XX
DE HIV-1 non-subtype B clone 94CY017-41 pol protein.
XX
KW HIV-1; human immunodeficiency virus; non-subtype B; gag; pol; env;
KW vpu; vif; vpr; tat; rev; nef; vaccine.
XX
OS Human immunodeficiency virus type 1.
XX
PN WO200026416-A1.
XX

```

```

PD 11-MAY-2000.
XX
XX 25-OCT-1999; 99WO-US24837.
XX
XX 02-NOV-1998; 98US-0184418.
XX
PA (UABR-) UAB RES FOUND.
XX
XX Hahn BH, Shaw GM, Gao F;
DR WPI; 2000-365651/31.
XX
XX Novel genomic nucleic acids of non-subtype B human immunodeficiency
PT virus type 1 useful for detecting and treating AIDS comprises a
PT specific nucleotide sequence
XX
XX Claim 41; Fig 15; 131pp; English.
XX
XX The present invention provides the protein and coding sequences for a
CC number of human immunodeficiency virus (HIV) type 1 non-subtype B
CC isolates. The sequences shown include the near full-length coding
CC sequences from each isolate, and the env, pol, vif, vpr, vpu, gag, tat,
CC rev and nef proteins. These can be used to detect the presence of HIV-1
CC in a sample and to produce antibodies against non-subtype B HIV-1 virus.
CC These antibodies can be used in vaccines to prevent and treat HIV
XX infection.
XX
XX Sequence 1003 AA;
XX
XX Query Match 57.5%; Score 3133; DB 21; Length 1003;
XX Best Local Similarity 58.3%; Pred. No. 9.6e-243;
XX Matches 575; Conservative 164; Mismatches 226; Indels 22; Gaps 6;
Qy 52 RAERET---IQSDRLGTAPRAG-GD--TIQATNGLAAPQFSLMKRPVVTAYIEGPV 105
Db 22 RANSPTSRELENGGRDNLPEAGTGQGTQISCN-----FPQITLWQRPVLTVKIEGQLK 76
Qy 106 EYLLDTGADDSIVAGIELGNYSPIKVGIGGIFNTKEYKNEIVEIKKVRATIMTGD 165
Db 77 EALLDTGADDTVLEELNLPCKWKPMIGGIGFVKVQYDQIAIEICGKRAIGTFLVGPT 136
Qy 166 PINIFGRNLTALGMSLNLPAKVEPIKIMLPGKDGPKLRQWPLTKKEIKALKEICEKM 225
Db 137 PYNIIIGRNMLVQLGCTLNPPIPIETVPVKLAPGMDGPKVKQWPLTEERIKALTEICEKM 196
Qy 226 EKEGOLLEAPPNPYNTPTFAIKKDKKWRMLIDPRELNKVTODTEIQLGIPHPAGLA 285
Db 197 EKEGKISIKGPNPNYNTPTFAIKKDKSTKRWLVDFRELNKRQDQFEVQLGIPHPAGL 256
Qy 286 KKRRTITVLDGDAYESIPLHEDFRPYTAFTLPSVNNAEPAKRYIKYKVLPGQWKGSPAF 345
Db 257 KKKAVTVLDGDAYFSVPLHEDFRKYTAFTIPSTNNETGVRVQYNVLPQGWKGSPAF 316
Qy 346 HTMRQVLEPRKANKDVIIIIQYMDLIIASDRTDLEHDRVVLQKELNGLGFGSTDPK 405
Db 317 SSMTKILEPFRSKNTELIYQYMDLVGSDLEISQHRVKIEELRAHLKKGFTYDPDKKH 376
Qy 406 OKDPPYHWMGYELWPTKWKLOKIQLOPEKEIWTNDIQKLVGLNWAQLYPGIKTKHL 465
Db 377 QKEPFLWNGYELHPDKWTVQPIKLPKDSVTNDIQKLVGLNWAQLYAGIKVKQLCK 436
Qy 466 LIIRKMTLIEEVQWTELAEELEENRIISQOECHYQOEKELEATVOKDOENQWYKH 525
Db 437 LLRGAKALTDIVTLTKAELEENREILKTPVHGYYDPSKDLIAEIQKQGDQWYQI 496
Qy 526 HQEE-KILKVGKYAKVKNTHNGIRLLAQVQKIGKEALVIGWGRIPKPHLPEREIVEQW 584
Db 497 YOEPFKNLTKGYAKRRSTHTNDIKQLTEAVQKIMTESIVIGKTPFKFLKIQKETWET 556
Qy 585 WDNVQWTVIPDWDVSTPPLVRLAFNLVGDPIPAETFTYDSCNRQSKGKAGYVTD 644
Db 557 WAQYQWATWIPWEFVNTPLVLKWLQLEKEPIAGAEYTVYVDAANRETKLKGAGYVTD 616

```



```

XX DR WPI: 2000-365651/31.
XX
PT Novel genomic nucleic acids of non-subtype B human immunodeficiency
PT virus type 1 useful for detecting and treating AIDS comprises a
PT specific nucleotide sequence -
XX
XX
PS Claim 41; Fig 15; 131pp; English.
XX
CC The present in invention provides the protein and coding sequences for a
CC number of human immunodeficiency virus (HIV) type 1 non-subtype B
CC isolates. The sequences shown include the near full-length coding
CC sequences from each isolate, and the env, pol, vif, vpr, vpu, gag, tat,
CC rev and nef proteins. These can be used to detect the presence of HIV-1
CC in a sample and to produce antibodies against non-subtype B HIV-1 virus.
CC These antibodies can be used in vaccines to prevent and treat HIV
CC infection.
XX
XX SQ Sequence 1002 AA;
Query Match 57.2%; Score 3120; DB 21; Length 1002;
Best Local Similarity 58.6%; Pred. No. 1.1e-241;
Matches 563; Conservative 169; Mismatches 222; Indels 6; Gaps 3;
QY 77 GATNRGLAA---PQSLWKRPVVTAYIEGQVPEVLLDTGADDSIVAGIELGNNYSPIKVG 133
Db 44 GAEGKGTISINLQITLWQRPVLTVRIGGOLIEALLDTGADDTVLEQINLPQKWKPMIG 103
QY 134 GIGGFINTKYEKVEIVLNKKVPRATIMTGTPTINIEGRNLTALGMSLNLPAKVEPIK 193
Db 104 GIGGFIVKQYDQILIEGKKAIGTVLVGPTPTINIEGRNLTALGMSLNLPAKVEPIK 163
QY 194 IMLKPGKDGKLPKOWPTEKIBALKEICEKMEKEGOLEBAPTPNYPNTPTFAIKKKDKN 253
Db 164 VKLPGIDGPKVQWPLTEKIKALTEICIDMEKEGKISGIPENPNTPIFALKKKDST 223
QY 254 KRWMLIDFRELNKVTDQTFEIQLGIPHPAGLAKRRITVLDVGDAFYSIPLHEDFRPYTA 313
Db 224 KRWKLVDFRELNRKTDQFWEVQLGIPHPAGLAKRRITVLDVGDAFYSIPLDENFRKYTA 283
QY 314 FTLPVNNAPGKRYIKVLPQGWKGSIPAIFQHTMRQVLEPFFRKANKDVIIQYMDILI 373
Db 284 FTIPSINNETPGIRYQYNVLPQGWKGSIPAIFQSMTKILEPFRTEPEIYQYMDLLYV 343
QY 374 ASDRTLEHDRVVLQKELLNGLGFSPTDEKFPKDPYHWMGYELWPTKWKLOKIQLPQK 433
Db 344 GSDETCQHRAKTEELRNHLLRGFTTPDKKHQKEPPLMWGELHDPDKWTVPQIQLPNK 403
QY 434 EIWTVNDIQKLVGLNMAQLYFGIKTKHLCLIRGKMTLTEEYQWTELAELAELENRII 493
Db 404 ESWTVNDIQKLVGLNMAQLYFGIKTKHLCLIRGKMTLTEEYQWTELAELAELENRI 463
QY 494 LSOEBOEGHYQEBKELEATVQKDOENQWTKIHQEE-KILKVGKAKVKKVTHNGIRLLA 552
Db 464 LKEPFGVGVYDPSKELTAELOKQCGDQWYIYQEPYKLNKTKGAKRGAHTNDVRKQLT 523
QY 553 QYVQKIGKALVIGRWPKPLPVERIWEQWMDNWOVTVIPDWDFVSTPPLVRLAFNL 612
Db 524 EAVQKATATESIVLWGVKPKELPIRKETWEVWVTEWQATWIPDWDFVSTPPLVRLAFNL 583
QY 613 VGDPIGAETFTYDGCNRSQKSGKAGYVTDGKDKVKKLEOTNQOAELEAFAMALDTS 672
Db 584 ETEPIPGAETFTYVVDGAANKETKLGAGYVTDGKDKVKKLEOTNQOAELEAFAMALDTS 643
QY 673 GPKVNIIVDSQYVWGLISASOPTSESKIVNQIIEEMIKKEAIYVAVWPAHKGIGGNOEVD 732
Db 644 GSEVNIIVDSQYVWGLISASOPTSESKIVNQIIEEMIKKEAIYVAVWPAHKGIGGNOEVD 703
QY 733 HLYSQIGTIOVLFLEKPEAQEHEKYSNVELSHKFGIPNLVARIQVNSCAOQOGEA 792
Db 704 KLYSSGIRKVLFDGDKAQEDHERHSHNWRAMASDFNLPIYAKELIVASCDCQLKGEA 763
QY 793 IHGQVNAELGTWQMDCTHLEGGKIIIVAVHVASGFIEAEVTPQESGROTALFLKLASRP 852

```

```

Db 764 MHGQVDCSPGIWQLDCTHLEGGKIIIVAVHVASGFIEAEVTPQESGROTALFLKLASRP 823
QY 853 ITHLHTDNGANFTSQEVKMWAWWIGIDEGSFGVPYNPQSGVVEAMNHLKKNQISRIEQA 912
Db 824 VKVHTDNGSNFTSAAMKACACWANIQQEFGIPYNPQSGVVEAMNHLKKNQISRIEQA 883
QY 913 NTIETIVLMAITHCMNFKRRGGIGDMTPSERLINMITTBOETQFLOAKNSKLKDFRVYFE 972
Db 884 EHLTAVQMAVFIINFRKGGIGGYSAGERIIDIIASDIQTKELQKQIKIQTIONFRVYRD 943
QY 973 GRDQLWKGPGELLKWKGEAVLVKVGTDIKIIPRRKAKIIRDYGRQEMDSGSHLEGARE 1032
Db 944 SRDPIWKGPAKLLKWKGEAVVQDNSEIKVVPVRRKVKIKIDYG--KOMAGDCVAGRQDE 1001

RESULT 24
AAP61508
ID AAP61508 standard; Protein; 1003 AA.
XX
AC AAP61508;
XX
DT 25-MAR-2003 (updated)
DT 19-AUG-1991 (first entry)
XX
DE Sequence of ARV-2 (9B) pol protein.
XX
KW LAV; HIV; ARV; HTLR; vaccine; AIDS; immunoassay; diagnosis;
KW lymphadenopathy syndrome.
XX
OS Human T-cell lymphotropic virus III.
XX
PN EPI81150-A.
XX
PD 14-MAY-1986.
XX
PF 30-OCT-1985; 85EP-0307860.
XX
PR 06-SEP-1985; 85US-0773447.
PR 31-OCT-1984; 84US-0667501.
PR 30-JAN-1985; 85US-0696534.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Luciw PA, Dina D, Steimer K, Pescador RS, Georgenasc C, Parkes D;
PI Hallewell R, Barr PJ;
XX
WP; 1986-126568/20.
DR N-PSDB; AAN60140.
XX
PT New recombinant human T-cell lymphotropic retro virus proteins -
PT useful in diagnostic immunoassays for antibodies in humans, and
PT in prodn. of monoclonal antibodies, as vaccines etc.
XX
PS Example; Fig 2; 67pp; English.
XX
CC The inventors claim a DNA construct contg. a DNA sequence
CC substantially as set forth in AAN60141, AAN60142, AAN60143, AAN60144,
CC which are each derived from AIDS-associated retroviruses. For the
CC purposes of this application, HTLV-III, LAV and ARV are generically
CC referred to as human T-cell lymphotropic retrovirus (HTLR). The
CC following recombinant polypeptides are also claimed: (a) ARV-2 p16 gag;
CC (b) ARV-2 p25 gag; (c) ARV-2 env; (d) ARV-2 p31 pol.
CC (Updated on 25-MAR-2003 to correct PA field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 1003 AA;
Query Match 57.2%; Score 3119; DB 7; Length 1003;
Best Local Similarity 56.8%; Pred. No. 1.3e-241;
Matches 567; Conservative 179; Mismatches 241; Indels 12; Gaps 6;
QY 38 GSTGEIYAAREKTERAERTIQ---GSDRGLTAPRAGDITIQTGATNRGLAAPQSLWKRP 94

```

```
Db 11 GKAREFSEQTRANSPTRRELQVWGGNNLS--EAGADR-QGTVSFNF--PQITLWQRP 65
Qy 95 VVTAYIEGQPVVLLDTGADDSIVAGIELGNYSKPIVGGIGGINTKEYKNVEIEVLNK 154
Db 66 LVTIRIGGOLKEALLDTGADDTVLEEMNLPQKWKPMIGGIGGIFKVRQYDQIPVEICGH 125
Qy 155 KVRATIMTGDPTINIFGRNLTALGMSLNLPAKVEPIKIMLPGKDGPKLRQWPLTKEK 214
Db 126 KAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLPQMGDPKVKQWPLTEK 185
Qy 215 IEALKEICKMEKEGOLEBAPTPNPTFAIKKKDKNKNRMLIDFRELNKVTDFTBEI 274
Db 186 IKALVEICTEMEKEGKISKIGPENPYTPVFAIKKDKSTKRWKLVDFRELNKRTQDFEV 245
Qy 275 QLGIPHAGLAKRRITVLVDGDAYESIPLHEDFRPYTAFTLPSVNNAPGPKRYIKVLP 334
Db 246 QLGIPHAGLKKKKSVITVLVDGDAYESVPLDKDFRYTAFTIPSNINNETPGRIQYNLP 305
Qy 335 QGKGSFAIFQHTMRQVLEPFPRKANKDVIIIOYMDLILASDRTDLEHDRVVLQKLELN 394
Db 306 QGKGSFAIFQSSMKILEPFRKQNPDIIVYOYMDLIVGSDLEIGQHTKTEELRQHLL 365
Qy 395 GLGFSTPDKFKDPPYHWMGELMPTKWLQKIQLPQKEIWTVDNDIQKLVGLNNAOQL 454
Db 366 RWGFTTPDKKHQKEPPFLMWGELHPDKWTQVQIMLPKSDSWTVNDIQKLVGLKNWASQI 425
Qy 455 YPGIKTKHLRLGRKMTLVEQWTELAELAELEENRIILSQEBGHYYQEEKELEATVQ 514
Db 426 YAGIKVKQCKLRLGRKALTEVPIPTEEAELEENREILKEPVHEVYDDPSKDLVAEIQ 485
Qy 515 KQENQWTKIHQEE--KILKVGKAKVKNTHNGIRLLAQVYQVKIKKEALVWGRIPKPH 573
Db 486 KOGQGWTVYIQEPPKNLKTKYARMGAHNDVKQLTEAVQKYSTESIVLWGIKPRK 545
Qy 574 LPVERIEQWMDNTWQVTWIPDWPVSTPPVLRLAFNLVGDPIPCAETFFYDGCNQRS 633
Db 546 LPIQKETWAWMXYQATWIPEWFEVNTPPVLKWLQYQLEKEPIVGAETFFYVDGAANRT 605
Qy 634 KEGKAGYVTRDKGKVKLEOTNQOAELEAFAMALTDSGPKVNIIVDSQVYMGISASOP 693
Db 606 KLKGAGYVTRDGRQKVVSTADTTNQKTELQALHLALQDSGLEVNIVTDSQYALGIIQAP 665
Qy 694 TESKIVNQIIEEMIKKEAIVAVWPAHKIGGNGQVDHLVSOGIROVLFLFLEKIEPAQE 753
Db 666 DKSELSVSOIIEQLIKKEKVLAWYPAHKIGGNGQVDKLVYAGIRKVLFLNGIDKAE 725
Qy 754 EHEKHSNVKELSHRFGIPNLVARQIVNSCAQCQKGEAIGHQVNAELGTWMDCTHLEG 813
Db 726 EHEKHSNWRWASDFNLPPVVAKEIVASCDCQLKGEAMHGQVDCSPGIWQLDCTHLEG 785
Qy 814 KIIIVAVHVASGFIEAEVIPSQSGROTALFLILKLASRPITHLHTDNGANFTSQEVKMYA 873
Db 786 KIILVAVHVASGYIEAEVIPAETGGETAFLKLAGRWPVKTIHTDNGSNFTSTIVKAAC 845
Qy 874 WWIGIEQSGFYPNPQSGQVBEAMNHLKNQISIREQANTTETIVLMAIHOVNPKRRGG 933
Db 846 WWAGIKQEFGIPNPQSGQVBEAMNHLKNQISIREQANTTETIVLMAIHOVNPKRRGG 905
Qy 934 IGDWTFSERLINWITEQEIQFLOAKNSKLDFRVYFREGROQLWKGGELLWKGEVAVL 993
Db 906 IGGYAGERIVDIIATDIQTKELQKITQIONFRVYVYRDNDKPLWKGPFAKLWKGEVAV 965
Qy 994 VRVGVDTIKIPRKAIRIDYGRQEMDSQSHLEGARD 1032
Db 966 IQDNSDKVVPVRKAKIINDYKQM---AGDDCVASROD 1001
```

RESULT 25

AAY77301

ID AAY77301 standard; Protein; 1003 AA.

XX

AC AAY77301;

```
XX 22-MAY-2000 (first entry)
XX HIV-1 (ATCC CRL 8597, single isolate) pol protein.
DE HIV-1; immunoassay; antigen; pol fragment; p31; recombinant; antibody;
KW detection.
XX Human immunodeficiency virus type 1 'ATCC CRL 8597'.
OS
XX
FH Key Location/Qualifiers
FT Misc-difference 1003
FT /note= "Encoded by GAT"
XX
PN US6013432-A.
XX
PD 11-JAN-2000.
XX
PF 17-MAY-1995; 95US-0443434.
XX
PR 08-JUL-1993; 93US-0089407.
PR 24-DEC-1987; 87US-0138894.
PR 17-AUG-1992; 92US-0931154.
PR 31-OCT-1984; 84US-0667501.
PR 30-JAN-1985; 85US-0696534.
PR 06-SEP-1985; 85US-0773447.
XX (CHIR ) CHIRON CORP.
XX
XX Luciw PA, Dina D;
PI
XX WPI; 2000-170256/15.
DR N-PSDB; AA290202.
XX
XX Immunoassay for antibodies against human immune deficiency virus, for
PT diagnosing infection, uses an immunogenic fragment of the pol protein
PT as antigen -
XX
PS Example 1; Fig 5B-E; 99pp; English.
XX
XX The invention relates to the improvement of HIV-1 immunoassays by the
CC use of an HIV-1 antigen comprising an immunogenic fragment of
CC recombinant or synthetic HIV-1 pol, which is encoded by an approximately
CC 9.7 kb sequence between a BstXI restriction site at position 3006 and an
CC NdeI site at position 5131 of the genome (the proviral DNA sequence is
CC given in AA290201). The immunogenic pol fragment is not immunologically
CC cross-reactive with human T cell lymphotropic viruses I or II. The
CC invention also encompasses the use of p31 as an antigen. The recombinant
CC antigens may be produced in Escherichia coli, Saccharomyces cerevisiae or
CC in mammalian cells. Immunoassays using the recombinant HIV proteins may
CC be used to diagnose and stage HIV-1 infections. Sequences
CC AAY77300-Y77302 represent, respectively, the gag, pol and env proteins
CC encoded by the proviral DNA of a single isolate of HIV-1 (ATCC CRL
CC 8597).
```

XX SQ Sequence 1003 AA;

```
Query Match 57.2%; Score 3119; DB 21; Length 1003;
Best Local Similarity 56.8%; Pred. No. 1.3e-241;
Matches 567; Conservative 179; Mismatches 241; Indels 12; Gaps 6;
Qy 38 GSTGIYAAAREKTERAERETIQ---GSDRGLTAPRAGGDTIQGATNRGLAAPQFSLWRRP 94
Db 11 GKAREFSEQTRANSPTRRELQVWGGNNLS--EAGADR-QGTVSFNF--PQITLWQRP 65
Qy 95 VVTAYIEGQPVVLLDTGADDSIVAGIELGNYSKPIVGGIGGINTKEYKNVEIEVLNK 154
Db 66 LVTIRIGGOLKEALLDTGADDTVLEEMNLPQKWKPMIGGIGGIFKVRQYDQIPVEICGH 125
Qy 155 KVRATIMTGDPTINIFGRNLTALGMSLNLPAKVEPIKIMLPGKDGPKLRQWPLTKEK 214
Db 126 KAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLPQMGDPKVKQWPLTEK 185
```

```
QY 215 IEALKEICEKMEKEGQLEAEAPPNTPTFAIKKDKKKNKMLIDFRELNKVTDQFTEI 274
Db 186 IKALVEICTEMEKEGKISKIGPENPTVFAIKKDKSTKWRKLVDFRELNKRTQDFWEV 245
QY 275 QLGIPHPAGLAKKRITVLDVGDAYSIFLHEDRPYTAFTPLPSVNNAEPCGRYIKVLP 334
Db 246 QLGIPHPAGLAKKRISVLDVGDAYSIFLHEDRPYTAFTPLPSVNNAEPCGRYIKVLP 305
QY 335 QGWGSPALFQHTMRQVLEPPRKANKDVIIQYMDLIIASDRTDLEHDRVVLQKLKLN 394
Db 306 QGWGSPALFQHTMRQVLEPPRKANKDVIIQYMDLIIASDRTDLEHDRVVLQKLKLN 365
QY 395 GLGSTPDEKQDPKPPHYHMGYELWPTKWKLOKIQLPQKEITWVNDIQKLVGNWAAQL 454
Db 366 RWGFTTDPKHKKEPPFLMWGYELWPTKWKLOKIQLPQKEITWVNDIQKLVGNWAAQL 425
QY 455 YPGIKTRHLCRLIGKMTLIEVQVWTELAELNRIILSQOECHYHQQEKELEATVQ 514
Db 426 YAGIKVKOLCKLRGKALTEVILPTEAELEAENREILKEPHEVYVDPKDLVARIQ 485
QY 515 KDOENOWTYKTHQEE-KILKVKYAKVKNTHNGIRLLAQQVQKIGKEALVIGRIKPFH 573
Db 486 KOGQOWTYQIYQPFNKLTKGYARMGAHTNDVKQTEAVQKVSTESIVWIKPIKPF 545
QY 574 LPVERIEQWMDNYQVWTPIDWDFVSTPLVRLAFNLVGDPIPGAETFTDGCNRSQS 633
Db 546 LPIQKTEWAEWMEYQATWIPWEPFVNTPLVQLVQLEKEPIVGAETFFVDGAANRET 605
QY 634 KEGAGYVTRDGKDKVKKLEQTTNOQAELEAFAMALTDSPKVINIVDSQVYMGISASQP 693
Db 606 KLGRAGYVTRDGKDKVKKLEQTTNOQAELEAFAMALTDSPKVINIVDSQVYMGISASQP 665
QY 694 TESEKIVNOIIEEMIKKEALYVAVWPAHKGIGGNOVDHLVSGIRQVLFLEKTEPAQE 753
Db 666 DKSESELVQIIEQLIKKEKYLAWPAHKGIGGNEQVDKLVSGIRKVLFLNGIDKAE 725
QY 754 EHEKHSNVKLSHKFGIPNLVARQIVNSCAQCOOGEAIIHQVNAELGTWQMDCTHLEG 813
Db 726 EHEKHSNWRMASDFNLPPVVAKEIVASCDKOLKGEAMHQVDCSPGIWQDCTHLEG 785
QY 814 KIIIVAVHVASGFEAEVPOESGROTALFLKLASBPITHLHTDNCANETSQEVKMA 873
Db 786 KIIIVAVHVASGFEAEVPOESGROTALFLKLASBPITHLHTDNCANETSQEVKMA 845
QY 874 WVIQESFGYPYNPQSGVVEAMNHLKNOISRQOANTIEIVLMAIHCNFKRRGG 933
Db 846 WVIQESFGYPYNPQSGVVEAMNHLKNOISRQOANTIEIVLMAIHCNFKRRGG 905
QY 934 IGMWTPSERLINMTTQEQIFLOAKNSKLDKFRVYFREGDQKWKGPGLLWKGEGAVL 993
Db 906 IGGYSAGERIVDIIATDITQTELQKITKIQNFRVYVRDNDKDLWKGPAKLLWKGEGAV 965
QY 994 VKVGTDIKIPRRKAKIIRDXGGQEMDSGSHLEGARD 1032
Db 966 IQDNSDIKVPRRKAKIIRDXGGQEMDSGSHLEGARD 1001
```

RESULT 26

AAY70601
ID AAY70601 standard; Protein; 1003 AA.

XX AAY70601;

XX 18-JUL-2000 (first entry)

DE Corrected version of wild type HIV pol protein.

XX HIV; gag; packaging cell line; lentivirus; retroviral vector particle;

KW pol; gene therapy; gene replacement; vaccine; biochemical reagent.

XX Human immunodeficiency virus.

XX WO200015819-A1.

```
XX 23-MAR-2000.
PD 10-SEP-1999; 99WO-US20675.
XX 11-SEP-1998; 98US-0100022.
PR 12-SEP-1998; 98US-0100063.
XX (CHTL-) CHILDRENS MEDICAL CENT.
PA Gray JT, Mulligan RC;
PI WPI; 2000-271455/23.
XX N-PSDB; AAZ52053.
PT New packaging cell line for producing a viral accessory protein
PT independent HIV derived retroviral vector particles, useful in gene
PT therapy or gene replacement -
XX Disclosure; Fig 9; 62pp; English.
XX The patent discloses new packaging cell line for producing a viral
CC accessory protein independent lentivirus, preferably
CC human immunodeficiency virus (HIV), derived retroviral vector particles.
CC The packaging cell line comprises a mammalian cell, a retroviral DNA
CC comprising a coding sequence for a lentivirus, preferably HIV, gagpol,
CC where the coding sequence has been mutagenised to improve expression of
CC the viral gagpol proteins, a second retroviral nucleotide sequence
CC comprising the coding sequence for a heterologous envelope protein and a
CC third retroviral nucleotide sequence comprising a DNA sequence of
CC interest and lentivirus, preferably HIV, cis-acting sequences required
CC for packaging, reverse transcription and integration.
CC The packaging cell lines and viral particles can be used for gene
CC therapy or gene replacement with improved safety. They can also be used
CC in the development and production of vaccines and biochemical reagents.
CC The present sequence is a corrected version of pol protein
CC from wild type HIV gagpol sequence.
CC The pol coding sequence is codon optimised to be used in the packaging
CC cell line. Codon optimisation results in improved expression of the
CC gagpol protein and reduces the risk of recombination between the transfer
CC vector and gagpol mRNA.
XX Sequence 1003 AA;
```

```
Query Match 57.2%; Score 3117; DB 21; Length 1003;
Best Local Similarity 56.9%; Pred. No. 1.9e-241;
Matches 570; Conservative 178; Mismatches 244; Indels 10; Gaps 6;
QY 32 PSGSSSGSTGEIYAAREKTERAERETIOGSDRGLTAPRAGGDTIOGATNRGLAAPQFSLW 91
Db 9 PQGKAREFSSTQTRANSPTREELQ--VWGRDNN-SLSEAGADR-QGTWS--FSFPQITLW 62
QY 92 KRPVVTAIEGQPVFVLDDTGADDSIVAGIELGNNYSPKIVGGIGGFINTKEYKNVEIEV 151
Db 63 QRLPVTIKIGQLKEALLDTGADDTVLEMLPGRWKPKMTGGIGGTFKVRQYDQILIEI 122
QY 152 LNKVRATIMTGDPTINIFGRNLTALGMSLNLPAKVEPIKMLKPGKDGPKLRQWPLT 211
Db 123 CGHKAIGTVIVGPTPVNIIGRNLTLQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLT 182
QY 212 KEKIEALKEICEKMEKEGQLEAEAPPNTPTFAIKKDKKKNKMLIDFRELNKVTDQF 271
Db 183 EEKIKALVEICTEMEKEGKISKIGPENPTVFAIKKDKSTKWRKLVDFRELNKRTQDF 242
QY 272 TEIOLGIPHPAGLAKKRITVLDVGDAYSIFLHEDRPYTAFTPLPSVNNAEPCGRYIK 331
Db 243 WEVOLGIPHPAGLAKKRISVLDVGDAYSIFLHEDRPYTAFTPLPSVNNAEPCGRYIK 302
QY 332 VLPQGWKSPALFQHTMRQVLEPPRKANKDVIIQYMDLIIASDRTDLEHDRVVLQKLK 391
Db 303 VLPQGWKSPALFQHTMRQVLEPPRKANKDVIIQYMDLIIASDRTDLEHDRVVLQKLK 362
QY 392 LLNGLGFSTPDEKQDPKPPHYHMGYELWPTKWKLOKIQLPQKEITWVNDIQKLVGNW 451
```

Db 363 HLLRWGFTTPDKKHQEPFLMWGYELHPDKWTVQPIVLPEKDSMTVNDIQKLVGLNNA 422
QY 452 AQLYPGIKTKHLRGLRGKMTLVEEYQWTELAELAELENRIILSQEGHYQEKELEA 511
Db 423 SIIYAGIKVRQCKLLRGKALTVEVPLTEEALEAENREILKEPVHGYVYDPSKDLIA 482
QY 512 TVOKDOENQWTKIHQEE-KILKVGKYAKVKNTHNGIRLLAQVQKIGKEALVINGRIP 570
Db 483 EIQKOGQGWTYQIYQEPFNKLTGKYARMKGAHTNDVKQLTEAVQKIATESIVWGKTP 542
QY 571 KFHLPVEREIQWQWNVQWVWIPDWDVFSVPLVRLAFNLVGDPIPGAETFYDGSN 630
Db 543 KEKLPQKETWEAWTEYQWATWPEWFEVNTPLVQLWYQLEKEPIGAETFYVDGAAN 602
QY 631 ROSKEKAGVYTDGRGKVKLLBQTNQOAELEAFAMALTDGSPKNNIIVDSQYVNGISA 690
Db 603 RETKLKAGVYTDGRGKVKVPLTDTNTQTELAQIHLAQDSGLEVINIVTDSQYALGIQ 662
QY 691 SQTSESKIVNOIIEEMIKKEAIYVAVPAHKIGGNQEVHLSQGIQVLFLEKIEP 750
Db 663 AOPDKSELSVQIIEQLIKKEKYLAWVPAHKIGGNQEVHLSQGIQVLFLEKIEP 722
QY 751 AQEEHEKYSNVKELSHKFGIPNLVARIQVNSCAQQCKGAEALHGQVNAELGTWQMDCTH 810
Db 723 AQEEHEKYSNVKELSHKFGIPNLVARIQVNSCAQQCKGAEALHGQVNAELGTWQMDCTH 782
QY 811 LEKIIIVAVHVASGPIEAETVPOESGROTALFLKLASRWPITLHTDNGANFTSOEVK 870
Db 783 LEKIVLAVHVASGPIEAETVPOESGROTALFLKLASRWPITLHTDNGANFTSOEVK 842
QY 871 MYAWTIGIEQSFVPYNPQSGVVEAMNHLNQSIRIREQANTIEIVLMAIHCNFKR 930
Db 843 AACWAGIKQEGIPYNPQSGVIESMKNELKKIIGQVQDAEHLKTAQMAVFIHFKR 902
QY 931 RGGIGMTSPERLINMITTEQELQFLOAKNSKLKDFRVYFREGRODLWKGPELWKGEG 990
Db 903 KGGIGYSAGERIVDIATDIQTKELQKQITQIONFRVYIRDSRDPVWKGPAKLLWKGEG 962
QY 991 AVLKVGTDIKIIPRKAIIIDYGGROEMDSGHLEGARE 1032
Db 963 AVVIQNSDIKVPRKAKIIRDYGKQM---AGDDCVASRQD 1001

RESULT 27
AA70602
ID AAY70602 standard; Protein; 1003 AA.
XX AAY70602;
DT 18-JUL-2000 (first entry)
XX AAY70602;
DE Codon optimised Human immunodeficiency virus pol protein.
XX HIV; gag; packaging cell line; lentivirus; retroviral vector particle;
KW pol; gene therapy; gene replacement; vaccine; biochemical reagent;
KW codon optimisation.
XX Human immunodeficiency virus.
OS Synthetic.
XX WO200015819-A1.
XX 23-MAR-2000.
XX 10-SEP-1999; 99WO-US20675.
XX 11-SEP-1998; 98US-0100022.
PR 12-SEP-1998; 98US-0100063.
XX (CHIL-) CHILDRENS MEDICAL CENT.
XX Gray JT, Mulligan RC;

XX WPI; 2000-271455/23.
DR N-PSDB; AA252054, AA252055.
XX New packaging cell line for producing a viral accessory protein
PT independent HIV derived retroviral vector particles, useful in gene
PT therapy or gene replacement -
XX Disclosure; Fig 9; 62pp; English.
XX The patent discloses new packaging cell line for producing a viral
CC accessory protein independent lentivirus, preferably
CC human immunodeficiency virus (HIV), derived retroviral vector particles.
CC The packaging cell line comprises a mammalian cell, a retroviral DNA
CC comprising a coding sequence for a lentivirus, preferably HIV, gagpol,
CC where the coding sequence has been mutagenised to improve expression of
CC the viral gagpol proteins, a second retroviral nucleotide sequence
CC comprising the coding sequence for a heterologous envelope protein and a
CC third retroviral nucleotide sequence comprising a DNA sequence of
CC interest and lentivirus, preferably HIV, cis-acting sequences required
CC for packaging, reverse transcription and integration.
CC The packaging cell lines and viral particles can be used for gene
CC therapy or gene replacement with improved safety. They can also be used
CC in the development and production of vaccines and biochemical reagents.
CC The present sequence is a pol protein from codon optimised HIV gagpol
CC sequence. The codon optimised gagpol sequence is used in the packaging
CC cell line. Codon optimisation results in improved expression of the
CC gagpol protein and reduces the risk of recombination between the transfer
CC vector and gagpol mRNA.
XX Sequence 1003 AA;
QY 32 PSGSSGSTGEIYAAREKTERAERETIQSGDRGLTAPRAGGTIOGATNRLAAPQSLW 91
Db 9 POGKAREFSSEQTRANSPTRRELQ--VMGRDNN-SLSEAGADR-QCTVS--FSFPQITLW 62
QY 92 KRPVVTAYTEGQVPEVLLDTGADDSIVAGIEIGNNYSKPIVGGIGGIFNTKEYKNVEIEV 151
Db 63 QRPLVTIKTGQQLKEALLDTGADDTVLEEMNLPGRWKPKMIGGIGFIVRQYDQDLIEI 122
QY 152 LNKVVRATIMTGDTPINIFGRNLTALGMSLNLPAKVEPIKIMLPGKDGPKLRWPLT 211
Db 123 CGHKAIGTVLVGFTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLPGMDGPKVQWPLT 182
QY 212 KEKIEALKBEICKEKEGOLEAPPTNPYNTPTFAIKKDKNKRMLIDFRELNKVTDQF 271
Db 183 EEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKDKSTKRRKLVDFRELNKRTQDF 242
QY 272 TEIQIGIHPAGLAKKRITVLDVGDAYSIPLHEDFRPYTAFTLPSVNNAPGKRYIK 331
Db 243 WEVQLGHPAGLAKKOKSVTVLDVGDAYSFVPLDKDFRKRYTAFTIPSINNPTPGIRYQN 302
QY 332 VLPOGWKGSPIAFQHTMROVLPFRKANKDVIIIOYMDILLIASDRDLEHVRVVLQKE 391
Db 303 VLPOGWKGSPIAFQCSMTKILEPFRKQNPDIYVQYMDLVSDEIQHRTKIELRQ 362
QY 392 LLNGLGFSTPDEKFDKPPYHMGYELMPTKWKLOKIQLPQKEIWTVDIQKLVGLNNA 451
Db 363 HLLRWGFTTPDKKHQEPFLMWGYELHPDKWTVQPIVLPEKDSMTVNDIQKLVGLNNA 422
QY 452 AQLYPGIKTKHLRGLRGKMTLVEEYQWTELAELAELENRIILSQEGHYQEKELEA 511
Db 423 SIIYAGIKVRQCKLLRGKALTVEVPLTEEALEAENREILKEPVHGYVYDPSKDLIA 482
QY 512 TVOKDOENQWTKIHQEE-KILKVGKYAKVKNTHNGIRLLAQVQKIGKEALVINGRIP 570
Db 483 EIQKOGQGWTYQIYQEPFNKLTGKYARMKGAHTNDVKQLTEAVQKIATESIVWGKTP 542
QY 571 KFHLPVEREIQWQWNVQWVWIPDWDVFSVPLVRLAFNLVGDPIPGAETFYDGSN 630

```

Db 543 KFKLPIQKTEWAWWTEWQATWPEWFEVNTPLVKLWQLEKEPIIGATFFYVDGAAN 602
Qy 631 ROSKEGAGYVDRGKDKVKKLEOTTNQOALEAFAMALTDSPKVNIIVDSQYVMGISA 690
Db 603 RETKLGRAGYVDRGQKVVPLTDTTNQKTELQAIHLALQDSGLEVNIVTDSQYALGIQ 662
Qy 691 SQPTSESKIVNQIIEEMIKKEALYVAVPAHKGIGGNOEVDHLVSOGIRQVLEKIEP 750
Db 663 AQDPKSESELVSLQIEQLIKKEKYLAWVPAHKGIGGNEQVDKLVSAIRKVLFDGDK 722
Qy 751 AQEEHEKHSNVKLSHKFGIPNLVARIQIVNSCAQCOQKGEA IHGQVNAELGTWQMDCTH 810
Db 723 AQEEHEKHSNWRAMASDENLPPVVAKEIVASCDKQLKGEAMHGQVDCSPGIWQLDCTH 782
Qy 811 LEGKIIIVAVHVASGFTEAEVIPAESGRQATFALLKLASRWPITHLTDNGANTSEV 870
Db 783 LEGKIIIVAVHVASGYTEAEVIPAETQATFALLKLASRWPITHLTDNGANTSEV 842
Qy 871 MVAMWIGIEQSFVPYNPQSGVVEAMNHLKNOISRIREQANTIEIVLMAIHCMMFKR 930
Db 843 AACWAGIKQEFPIPNPQSGVIESMKNELKLIQGVROQAHLKTAQVMAVFIHFKR 902
Qy 931 RGGIGDMPSERLINMITTQEIQFLOAKNSKLKDFRYVYFREGDQLWKGPGELLWKGE 990
Db 903 KGGIGGYSAGERIVDIATDIQTKELQKITKIQNFRVYRDSRDPVWKGPAKLLWKGE 962
Qy 991 AVLKVGKTDKIIPRRKAKIIRDYGGROEMDSGSHLEGARE 1032
Db 963 AVVIQNSDIKVVPRKAKIIRDYGKOM---AGDDCVASRQD 1001

```

RESULT 28

AAB69284 standard; Protein; 1003 AA.

XX AC AAB69284;

XX DT 20-APR-2001 (first entry)

XX DE HIV-1 non-subtype B clone 93BR029-4 pol protein.

XX KW HIV-1; human immunodeficiency virus; non-subtype B; gag; pol; env;

XX KW vpu; vif; vpr; tat; rev; nef; vaccine.

XX OS Human immunodeficiency virus type 1.

XX PN W0200026416-A1.

XX PD 11-MAY-2000.

XX PF 25-OCT-1999; 99WO-US24837.

XX PR 02-NOV-1998; 98US-0184418.

XX PA (UABR-) UAB RES FOUND.

XX PI Hahn BH, Shaw GM, Gao F;

XX DR WPI; 2000-365651/31.

XX Novel genomic nucleic acids of non-subtype B human immunodeficiency

XX virus type 1 useful for detecting and treating AIDS comprises a

XX specific nucleotide sequence -

XX Claim 41; Fig 15; 131pp; English.

XX The present invention provides the protein and coding sequences for a
 CC number of human immunodeficiency virus (HIV) type 1 non-subtype B
 CC isolates. The sequences shown include the near full-length coding
 CC sequences from each isolate, and the env, pol, vif, vpr, vpu, gag, tat,
 CC rev and nef proteins. These can be used to detect the presence of HIV-1
 CC in a sample and to produce antibodies against non-subtype B HIV-1 virus.

CC These antibodies can be used in vaccines to prevent and treat HIV
 CC infection.

XX SQ Sequence 1003 AA;

Query Match 57.2%; Score 3116.5; DB 21; Length 1003;
 Best Local Similarity 56.9%; Pred. No. 2e-241;
 Matches 570; Conservative 173; Mismatches 250; Indels 9; Gaps 4;

```

Qy 32 PSGSSSGSTGEIYAAREKTERAERETIQGSDRGUTAPRAGGDTTQAGATNRGLAAPQFSLW 91
Db 9 PQGKAREFPSSQTRANSPTS---ELQVWGRGNSLSSETGADRGQDVSFGF--PQITLW 62
Qy 92 KRPVVTAYIEQPVVEVLLDTCADDSDIAGIELGNNSPKIVGGTGGFTNTKEYKNVELEV 151
Db 63 QRLVTVKIGQLKEALDTCADDTVLEENLPGRWKPMGIGGFKVQYDQIPIEI 122
Qy 152 LNKVRATITMGDTPINIFGRNITALTGMSLNLVAKVEPIKIMLPGKDGPKLRQWPLT 211
Db 123 CGRKATGVLVGPVPIVNIIGNLLTQIGCTLNFPIETVPIVVKLPGMDGPRVKQWPLT 182
Qy 212 KEKTEALKEICEKMEKEGQLEEAAPTNPYNTPTFAIKKKDKKNKRWMLIDFRELKNVQDF 271
Db 183 EEKIKALTEICTEMEKEGKISKIGPENPNTPVFAIKKKDKSTKWRKLVDFRELKRTQDF 242
Qy 272 TEIOLGIPHPAGLAKKRITVLDVGDVAFVSTPLHEDPRPYTAFTLPSVNNAEPCGRKIYK 331
Db 243 WEVOLGIPHPAGLAKKRSVTVLDGDVAFVSPLDKFRKYTAFTPTNNPTGRLUYQIN 302
Qy 332 VLPQGWKGSPIAFQHTMRQVLEPFERKANKDVIILQYMDDIILIASDRTDLEHDRVVLQKE 391
Db 303 VLPQGWKGSPIAFQSSMTKILEPFRKQNPDIYIQYMDLVGSDLEIGHQRTKEELRQ 362
Qy 392 LLNGLGFSTDEKQDPPYHWMGYELMPTTKWLQKIQLPQKEIWTWVNDIQKLVGLNWA 451
Db 363 HLLRWGFTTDPDKHKQKPPFLWMGYELMPTVQPIVLPEKDSWTWVNDIQKLVGLNWA 422
Qy 452 AOLYPGIKTHLCRLIRGKMTLTEEVQWTELAELLENRIILSQEGHGYEKEKELEA 511
Db 423 SQIYAGIKVROLCKLKGTKALTEVPLTAPAELEAENRELLKPEVHGVYDYSKDLIA 482
Qy 512 TVQKQDENQWTKIHQEE-KILKVGKYAKVKNTHNGIRLAAQVQVQKIGKEALVWGRIP 570
Db 483 EIQKQGGQWYQYIYQEPYKLNKTKGYARMRGATNDVKQLTEAVQKITTESIVWKGIP 542
Qy 571 KFHLPVEREWEQWMDNMQVWTIPDWDVFPPLVRLAFNLVGDPIPGATFTYDSCN 630
Db 543 KFKLPIQKTEWAWWIEYQATWPEWFEVNTPLVKLWQLEKEPIVGAETFFYVDGAAN 602
Qy 631 ROSKEGAGYVDRGKDKVKKLEOTTNQOALEAFAMALTDSPKVNIIVDSQYVMGISA 690
Db 603 RETKLGRAGYVDRGQKVVPLTDTTNQKTELQAIHLALQDSGLEVNIVTDSQYALGIQ 662
Qy 691 SQPTSESKIVNQIIEEMIKKEALYVAVPAHKGIGGNOEVDHLVSOGIRQVLEKIEP 750
Db 663 AQDPKSESELVSLQIEQLIKKEKYLAWVPAHKGIGGNEQVDKLVSAIRKVLFDGDK 722
Qy 751 AQEEHEKHSNVKLSHKFGIPNLVARIQIVNSCAQCOQKGEA IHGQVNAELGTWQMDCTH 810
Db 723 AQEEHEKHSNWRAMASDENLPPVVAKEIVASCDKQLKGEAMHGQVDCSPGIWQLDCTH 782
Qy 811 LEGKIIIVAVHVASGFTEAEVIPAESGRQATFALLKLASRWPITHLTDNGANTSEV 870
Db 783 LEGKIIIVAVHVASGYTEAEVIPAETQATFALLKLASRWPITHLTDNGANTSEV 842
Qy 871 MVAMWIGIEQSFVPYNPQSGVVEAMNHLKNOISRIREQANTIEIVLMAIHCMMFKR 930
Db 843 AACWAGIKQEFPIPNPQSGVIESMKNELKLIQGVROQAHLKTAQVMAVFIHFKR 902
Qy 931 RGGIGDMPSERLINMITTQEIQFLOAKNSKLKDFRYVYFREGDQLWKGPGELLWKGE 990
Db 903 KGGIGGYSAGERIVDIATDIQTKELQKITKIQNFRVYRDSRDPVWKGPAKLLWKGE 962

```

QY	991	AVLVKVGTDIKILPRRKAIIIRDYGROEMDSGSHLEGARED	1032
Dd	963	AVVIQNDSIDIKVVPKRKKIIRDYG--KOMAGGDCVAGRQDE	1002
 RESULT 29 AAR29705			
ID	AAR29705	standard; Protein; 1003 AA.	
XX	AC		
XX	AC		
DT	25-MAR-2003	(updated)	
DT	26-APR-1993	(first entry)	
XX			
DE	pol gene decoded from viral DNA from hTLR.		
XX			
KW	express proteins: lymphadenopathy syndrome; AIDS; HIV; HTLV;		
KW	serological immunoassays; antibodies to hTLR; monoclonal antibodies;		
KW	probes; ss.		
OS	Human T cell lymphotropic retrovirus.		
XX			
PX	EP518443-A1.		
PN			
PD	16-DEC-1992..		
XX			
PF	30-OCT-1985; 92EP-0201711.		
XX			
PR	31-OCT-1984; 84US-0667501.		
PR	30-JAN-1985; 85US-0696534.		
DR	06-SEP-1985; 85US-0773447.		
XX			
PA	(CHIR) CHIRON CORP.		
XX			
PI	Barr PJ, Dina D, George-Nascimento C, Hallewell R;		
PI	Luciw PA, Parkes D, Pescador RS, Steimer K, Truett M;		
XX			
DR	WPI; 1992-417329/51.		
DR	N-PSDB; AAO31936.		
XX			
PT	Recombinant DNA construct including replication system recognised		
PT	by unicellular microorganism - used to form recombinant proteins		
PT	for diagnosing AIDS and lymphadenopathy syndrome		
XX			
PS	Example 5; Fig 2; 32pp; English.		
XX			
CC	This sequence was decoded from the pol gene from proviral hTLR DNA.		
CC	Proteins associated with lymphadenopathy syndrome and/or AIDS may		
CC	be used in serological immuno-assays to detect antibodies to hTLR.		
CC	The polypeptides can be used alone or in fusion constructs to		
CC	produce antisera or monoclonal antibodies which may be used for		
CC	therapy or diagnosis.		
CC	(Updated on 25-MAR-2003 to correct PN field.)		
CC	(Updated on 25-MAR-2003 to correct PF field.)		
XX			
SQ	Sequence 1003 AA;		
 Query Match 57.2%; Score 3116; DB 13; Length 1003; Best Local Similarity 56.7%; Pred. No. 2.2e-241; Matches 566; Conservative 180; Mismatches 241; Indels 12; Gaps			
QY	38	GSTGIYAAREKTERAERETIQ---GSDRGITAFRAPGGDIQCATNRLAAPGFSLWKRP	94
Dd	11	GKAREFSSEQTRANSPTRELOVWGGENSLS--EAGADR-QGTVSFNF--POITLWRQP	65
QY	95	VVTAYIEQGPPVEVLIDTGADDSIVAGIELGNYSFKIVGGIGGFINTEKEYKNVEIEVLNK	154
Dd	66	LVTIRIGQLKEALLDTGADDTVLEENLLPKCKMKMIGSGFIKVROYDQIPVEICGH	125
QY	155	KVRATIWTGDTPPINIFGRNILTALGMSLNLPVAKVEPIKTMLKPDKDGPKLRQWPJTKEK	214
Dd	126	RAIGVPLVGPFPVNIIGNLTQTGCFTNPISPIETPVKLPGMDGPKVKOWPLEEK	185

CC A fragment of DNA approximately 200-500 base pairs in length is
CC ligated into a vector and used to transform E.coli. These cells then
CC express a polypeptide which is immunoreactive with HTLV-III-specific
CC antibody. The HTLV-III polypeptides can be used for the production
CC of antibodies, in immunosays for the detection of HTLV-
CC III-specific antibodies and in vaccines for the prevention of AIDS.
CC The antibodies can also be used to detect HTLV-III polypeptides.
CC (Updated on 25-MAR-2003 to correct PN field.)
CC (Updated on 25-MAR-2003 to correct PF field.)
CC (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 1015 AA;

Query Match 57.1%; Score 3114; DB 14; Length 1015;
Best Local Similarity 56.4%; Pred. No. 3.3e-241;
Matches 566; Conservative 177; Mismatches 241; Indels 20; Gaps 4;

QY 46 ARE---KTERAERTTQSGDRGLTAPRAGGDTTQAGATNRGLAA-----PQFS 89
DB 13 AREFSSEQTRANSPTTISSEQTRANSPTRELQVWGRDNLSSEAGADRGCTVSNFPQIT 72

QY 90 LAKRPVVTAYIEQOPVEVLDTGADDSIVAGIELGNYSKPIVGGIGGINTKEYNVEI 149
DB 73 LMORPLVTIKIGQLKEALLDTGADTVLEEMSLPGRWKPKMIGGIGGIFKVRQYDQILI 132

QY 150 EVLNKKVRATIMTGDPTPIFNIGNLTALGMSLNLPAKVEPIKIMLKPGKDGPKLRQWP 209
DB 133 EICGHKAIGTVLGPVPIVNIIGNLLTQIGCTLNFPISPIETVPVCLKPGMDGPKVQWP 192

QY 210 LTKEKEALKEICEKMEKQGLQEEAPPTNPPTFAIKKKKKNKWRMLIDFRELNVQ 269
DB 193 LTEEKIKALVIEITEMEKEGKISIGPENPNTPVFAIKKKDKSTKRWKLVDRELNRRTQ 252

QY 270 DFEIOLGPHPAAGLAKKRITVLDVGDAYFESPLHEDRPYTAFTLPSVNAEPCKRI 329
DB 253 DFWEVLQGLPHPAAGLAKKRSVTVLDVGDAYFVPLDEDFKRYTAFTIPSNNETPGSG 312

QY 330 KYVLPGQWKGSPAIFQTRQVLEPFRKANKDVIITQYMDLILIASDRTDLEHDRVVLQ 389
DB 313 YNVLPGQWKGSPAIFQSSMTKILEPFRKQNPDIIVQYMDLVGSDLEIGQHRKIEEL 372

QY 390 KELLNGLSTPDEKQKQPPHWMGYELWPKWKLOKLOLPOKETWTVNDQKLVGN 449
DB 373 ROHLRWGFTTPDKKQKQEPPLFMWGYELHPDKWTQIPLPEKDSWTVNDQKLVGN 432

QY 450 WAAOLYPGKTKHLCLRLIRKMTLVEEQWTELAELEENRILTSQEQEGHYQEKE 509
DB 433 WASQIIPGKIKVROLCKLRGTALTEVILPTEAELELEAENRILKEPVGHYIDPSKOL 492

QY 510 EATVQKQENQWTKYIHOEE-KILKGYKAKVKNTHNGIRLLAQVYQVQKIGKEALVWGR 568
DB 493 IAEIQOGOGQWYQIYQEPFNKLTGKYARMGAHTNDVKOLTEAVQKITTESIIVWCK 552

QY 569 IPKFHLPLVERIEWEQWMDYQVWTWIPDNDVSTPPLVRLAFNLVGDPIPGAEFTYDGS 628
DB 553 TPKEFLPIQKETWETWTWTEQATWIPEWFEFNTPLVWLWYOLEKEPIVGAFTFYVDGA 612

QY 629 CNRQSEKAGYVTDGKDKVKLEOTNTQOAELEAFALTDGSPKVNIIYDSQVVMGI 688
DB 613 ASRETKLAGYVTVNGROKQVYLTHTTQKTELQAINALQDSGLEVNIYDSQVYALGI 672

QY 689 SASQPTSESKIVNQIIEEMIKKAEIYVAVWPAHKGGIGGQNEVDHLVSGIGRQVLFLEKI 748
DB 673 IQAQPDKSESELVNIQIEILIKKEKYLAWVPAHKGGIGGQNEVDKLVSAIGIRKILFDGI 732

QY 749 EPAQEEHEKXVSHVSKLSHFKGIPNLVARQIVNSCAQCOQKGEAIGHQVNAELGTWQMD 808
DB 733 DKAQEEHEKXVSHVNAWASDFNLPPVVAKEIIVASCDKQLKGEAMHGQVDCSPGIIQLDC 792

QY 809 THLEGRKIIIVAVHVASGFTAEAVIPOESGRQFALFLLKLSRWPIPTHLTNDGANTSOE 868
DB 793 THLEGRKIIIVAVHVASGFTAEAVIPEATQETAYFLLKLAGRPVYKTIHTDNGSNFTSAT 852

QY 869 VKMAVWVIGTIEQSFQVYPNPQSQGVVEAMNHHHLKNOISRIREQANTTETIVLMAIHC MNF 928
DB 853 VRAACWAGIKQEFGIPYNPQSQGVVESNMKELKTIQGVQDAEHLKTAVQMAVEITHNF 912

QY 929 KRRGGTMDTPSERLINMTTQETQFLOAKNSKLDFRYVYFREGRDQLWKGCGELLWKG 988
DB 913 KRRGGIGGYSAGERIVDIATDIOTKELQKITKIONFRVYVYRDSRNLWKGPAKLLWKG 972

QY 989 EGAVLVKVTGTDKIIPRRKAKIIRDYGGRRQEMDSGSHLEGARE 1032
DB 973 EGAVVIQDNSDIKVPVRRKAKIIRDYGKQ---AGDDCVASROD 1013

RESULT 32
AAP61507
ID AAP61507 standard; Protein; 1012 AA.
XX AAP61507;
XX XX
DT 08-JUN-1991 (first entry)
XX Sequence of reverse transcription ("pol").
XX HIV; LAV; AIDS; diagnosis; vaccine.
XX HTLV-IIIIB/H9 cells (ATCC CRL 8543).
XX EP187041-A.
XX 09-JUL-1986.
XX 23-DEC-1985; 85EP-0309454.
PR 24-DEC-1984; 84US-0685272.
PR 04-DEC-1985; 85US-0805069.
XX (GETH) GENENTECH INC.
XX Capon DJ, Lasky LA;
XX WPI; 1986-177602/28.
DR N-PSDB; AAN60288.
XX Acquired immune deficiency syndrome polypeptide(s) - obt'd. by
PT molecular cloning etc. and used for diagnosis and in vaccines
PT against virus disease
XX Example; fig 2; 125pp; English.
XX A comparison of AAN60287 with the cDNA of the HTLV-III genome
CC revealed one particular clone, designated p7.11 which contained a
CC DNA sequence encoding this peptide (AAP60308) sequence. This approx.
CC 2.2 kilobase covers the precursor gag region and encodes, 5' to 3',
CC p-12, p-15, p-24 a second p-15 protein, and approx. 300 extra base
CC pairs 3' to the gag region (see AAN60288).
XX Sequence 1012 AA;

Query Match 57.1%; Score 3113; DB 7; Length 1012;
Best Local Similarity 56.2%; Pred. No. 4e-241;
Matches 563; Conservative 183; Mismatches 249; Indels 6; Gaps 3;

QY 33 SGSSSGSTGEIYAAREKTERAERTTQSGDRGLTAPRAGGDTTQAGATNRGLAAPQFS 92
DB 15 SEQTRANSPTTISSEQTRANSPTRELQVWGRDNLSSEAGADRGCTVSNF---PQITLWQ 72

QY 93 RPVVVTAYIEQOPVEVLDTGADDSIVAGIELGNYSKPIVGGIGGINTKEYNVEI 152
DB 73 RPLVTIKIGQLKEALLDTGADTVLEEMSLPGRWKPKMIGGIGGIFKVRQYDQILIEIC 132

QY 153 NKKVRATIMTGDPTPIFNIGNLTALGMSLNLPAKVEPIKIMLKPGKDGPKLRQWPLTK 212
DB 133 GHKAIGTVLGPVPIVNIIGNLLTQIGCTLNFPISPIETVPVCLKPGMDGPKVQWPLTE 192

PN	US5851813-A.
XX	
PD	22-DEC-1998.
XX	
PF	27-JAN-1994; 94US-0188583.
XX	
PR	27-JAN-1994; 94US-0188583.
PR	12-JUL-1990; 90US-0551945.
PR	09-JUL-1991; 91US-0727494.
XX	
PA	(HARD) HARVARD COLLEGE.
XX	
PI	Desrosiers RC;
XX	
DR	WPI; 1999-080408/07.
DR	N-PSDB; AAV81866.
XX	
PT	Lentivirus antigenic compositions - containing lentivirus with nef gene deletion
PT	
XX	
PS	Disclosure; Fig 2A-R; 93pp; English.
XX	
CC	The invention relates to an antigenic composition comprising an isolated primate lentivirus whose genome contains an engineered non-revertible null mutation in the nef gene, or an infectious DNA clone in a carrier.
CC	The antigenic composition is used in vaccines against infection by the lentivirus, e.g. AIDS.
CC	
XX	Sequence 1012 AA;
SQ	
	Query Match 57.1%; Score 3113; DB 20; Length 1012;
	Best Local Similarity 56.2%; Pred. No. 4e-241;
	Matches 563; Conservative 183; Mismatches 249; Indels 6; Gaps 3;
QY	33 SGSSSGTGEIYAAREKTERAERETIQSDRGLTAPRAGGDTIQATNRGLAAPQFSLWK 92
Db	15 SEQTRANSPITISSEQTRANSPTRRELQVWGRDNNSPSEAGADROGTVSFNF--PQITLWQ 72
QY	93 RPVVTAYTEGQPVVLLDTGADDSTVAGIELGNYSKPIVGGIGGIFINTKEVKNVEIYL 152
Db	73 RPLVTIKGGQKKEALLDTGADDVLEEMSLPGRWKPMKIGGIGFIVKQIDQLILIEIC 132
QY	153 NKKVRATIMTGDTPINIFGRNLITAGLSNLPAKAVEPIKIMLPKGDGPKLRWPLTK 212
Db	133 GHKALGTVLVGPVPTVNIIGRNLLQIGCTLNFPISTPIETVPVKLPKGMGPKVKQWPLTE 192
QY	213 EKIEALKEICEKMEKEGQLEAEAPPNTPTFAIKKDKNKKRMLIDFRELNKVTDFT 272
Db	193 EKIKALVEICTEMEKEGKISKIGPENPTVPFAIKKDKSTKWKLVDFRELNKRTQDFW 252
QY	273 EIQGIPHPAGLAKRRITVLDVGDAFSPILHEDFRPYTAFTLPSVNNAPGKRYIKV 332
Db	253 EVQIGIPHPAGLKKKSVTVLDVGDAFSPVLPDEDEFRKYTAFTTIPSINNETPGIRYQNV 312
QY	333 LPOGKSGPAIFQHTMROVLPPFRKANKDVIIIOYMDLIIASDRTDLEHDRVVLQKEL 392
Db	313 LPOGKSGPAIFQSSMTKILEPFRKQNDVIYQYMDLIVGSDLEIGQHRKTEELRQH 372
QY	393 LNGLGFSTPDEKFOKPPYHMGYELWPTKWKLOKIQLPQKEIYTVNDIQKLVGLNAA 452
Db	373 LLRWGLTTPDKKHQKEPFLMNGYELHPDKWTQVPLPEKDSWTVNDIQKLVGLNAA 432
QY	453 QLYPGIKTKHLCLRLIRGKMTLVEEVOWTELAELAEENRIILSOEQEGHYOEKELEAT 512
Db	433 QIYPGKIKVQCKLLRGTKALTVEIPLTEEALELAENREILKEPVHGVDPSKOLIAE 492
QY	513 VQKDOENQWTKYIHQEE-KILKVGKYAKVKNHTNGIRLLAQVQKIGKEALVWGRIPK 571
Db	493 IQKQGQGWTVQIYQEPFNKLTGKYARMGAHTNDVKQLTEAVQKITTESIIVGWKTPK 552
QY	572 FHLPEREINQWMDNYQWTVIPDWDFVSTPPLVRLAFNLVGDPIPGAETFFYDGSNCR 631
Db	553 FKLPIQKETWETWTEYQWATWPEWFEVNTPPVLKWLWYQLEKEPIVGAETFFYDGAANR 612
QY	632 QSKGKAGYVTRGDKVKKLEQNTNOQAELEAFAMALTDGPKVNIIVDSQYVNGISAS 691
Db	613 ETRLGKAGYVTRGKQKVVPLTNTNQKTELQAIYALQDSGLEVNIIVTDSQYALGIQOA 672
QY	692 OPTESKIVNQIIEEMIKKAIYVAVPAHKGIGGNOVDHLSVSGIROVLEKIEIPA 751
Db	673 QPDQSESELVNIQIIEQLKKEVYLAWPAPHKGIGGNOVDKLVASGIRKILFLDGDKA 732
QY	752 QEEHEKYSNKKELSHKFGIPNLVARQIVNSCAQOQKGEAIGHQVNAELGTWQMDCTHL 811
Db	733 QDEHEKYSNKRAMASDFNLPPVVAKEIVASCDCQLKGEAMHGQVDCSPGIWQLDCTHL 792
QY	812 EKKIIVAVHVASGIEAEVTPQESGROQTALFLKLASRWPITHLHTDNGANFTSQEVKM 871
Db	793 EKGVLVAVHVASGIEAEVTPQESGROQTALFLKLASRWPVKT IHTDNGSNFTSATVKA 852
QY	872 VAWWIGIEQSGVPYNQSQGVVAMNHLKNOISRIREQANTTETIVLMAIHCNFKRR 931
Db	853 ACWAGIKQEGIPYNQSQGVVESMKNELKIIGVROQAEHLKTAQVAVFIHNEFKR 912
QY	932 GGIGDMTPSERLINNITTEQIEQFLOAKNSKLKDFRVYFREGROQLWKGPGLLWKGGA 991
Db	913 GGIGYSAGERIVDIIATDIQTKELQKITKIONFRVYRDSRNPMLKGPAPKLLWKGBGA 972
QY	992 VLVKVGVDIKIIPRKKAKIIRYDGGREMDSGSHLEGARED 1032
Db	973 VVIQDNDSIKVVPVRKAKIIRYDGKQ---AGDDCVASROD 1010
RESULT 33	
ID	AAW89323
XX	
AC	AAW89323 standard; Protein; 1012 AA.
XX	
DT	01-JUN-1999 (first entry)
XX	
DE	HIV-1 pol protein sequence.
XX	
KW	Antigenic composition; primate; lentivirus; nef gene; vaccine;
XX	
OS	infection; AIDS; HIV-1; pol protein.
XX	
OS	Human immunodeficiency virus type 1.
XX	
PN	US5851813-A.
XX	
PD	22-DEC-1998.
XX	
PF	27-JAN-1994; 94US-0188583.
XX	
PR	27-JAN-1994; 94US-0188583.
PR	12-JUL-1990; 90US-0551945.
PR	09-JUL-1991; 91US-0727494.
XX	
PA	(HARD) HARVARD COLLEGE.
XX	
PI	Desrosiers RC;
XX	
DR	WPI; 1999-080408/07.
DR	N-PSDB; AAV81866.
XX	
PT	Lentivirus antigenic compositions - containing lentivirus with nef gene deletion
PT	
XX	
PS	Disclosure; Fig 2A-R; 93pp; English.
XX	
CC	The invention relates to an antigenic composition comprising an isolated primate lentivirus whose genome contains an engineered non-revertible null mutation in the nef gene, or an infectious DNA clone in a carrier.
CC	The antigenic composition is used in vaccines against infection by the lentivirus, e.g. AIDS.
CC	
XX	Sequence 1012 AA;
SQ	
	Query Match 57.1%; Score 3113; DB 20; Length 1012;
	Best Local Similarity 56.2%; Pred. No. 4e-241;
	Matches 563; Conservative 183; Mismatches 249; Indels 6; Gaps 3;
QY	33 SGSSSGTGEIYAAREKTERAERETIQSDRGLTAPRAGGDTIQATNRGLAAPQFSLWK 92
Db	15 SEQTRANSPITISSEQTRANSPTRRELQVWGRDNNSPSEAGADROGTVSFNF--PQITLWQ 72
QY	93 RPVVTAYTEGQPVVLLDTGADDSTVAGIELGNYSKPIVGGIGGIFINTKEVKNVEIYL 152
Db	73 RPLVTIKGGQKKEALLDTGADDVLEEMSLPGRWKPMKIGGIGFIVKQIDQLILIEIC 132
QY	153 NKKVRATIMTGDTPINIFGRNLITAGLSNLPAKAVEPIKIMLPKGDGPKLRWPLTK 212
Db	133 GHKALGTVLVGPVPTVNIIGRNLLQIGCTLNFPISTPIETVPVKLPKGMGPKVKQWPLTE 192
QY	213 EKIEALKEICEKMEKEGQLEAEAPPNTPTFAIKKDKNKKRMLIDFRELNKVTDFT 272
Db	193 EKIKALVEICTEMEKEGKISKIGPENPTVPFAIKKDKSTKWKLVDFRELNKRTQDFW 252
QY	273 EIQGIPHPAGLAKRRITVLDVGDAFSPILHEDFRPYTAFTLPSVNNAPGKRYIKV 332
Db	253 EVQIGIPHPAGLKKKSVTVLDVGDAFSPVLPDEDEFRKYTAFTTIPSINNETPGIRYQNV 312
QY	333 LPOGKSGPAIFQHTMROVLPPFRKANKDVIIIOYMDLIIASDRTDLEHDRVVLQKEL 392
Db	313 LPOGKSGPAIFQSSMTKILEPFRKQNDVIYQYMDLIVGSDLEIGQHRKTEELRQH 372
QY	393 LNGLGFSTPDEKFOKPPYHMGYELWPTKWKLOKIQLPQKEIYTVNDIQKLVGLNAA 452
Db	373 LLRWGLTTPDKKHQKEPFLMNGYELHPDKWTQVPLPEKDSWTVNDIQKLVGLNAA 432
QY	453 QLYPGIKTKHLCLRLIRGKMTLVEEVOWTELAELAEENRIILSOEQEGHYOEKELEAT 512
Db	433 QIYPGKIKVQCKLLRGTKALTVEIPLTEEALELAENREILKEPVHGVDPSKOLIAE 492
QY	513 VQKDOENQWTKYIHQEE-KILKVGKYAKVKNHTNGIRLLAQVQKIGKEALVWGRIPK 571
Db	493 IQKQGQGWTVQIYQEPFNKLTGKYARMGAHTNDVKQLTEAVQKITTESIIVGWKTPK 552
QY	572 FHLPEREINQWMDNYQWTVIPDWDFVSTPPLVRLAFNLVGDPIPGAETFFYDGSNCR 631
Db	553 FKLPIQKETWETWTEYQWATWPEWFEVNTPPVLKWLWYQLEKEPIVGAETFFYDGAANR 612

[illegible]

Qy	92	KRPVVYAYIEGQPVVEVLDTGADDISIVAGIELGNYSKPIVGGIGGFINTEKYKNVEIEV	151
Db	63	QRPLVTKIGGQLEALDTCADDTVPLEEMNLPGRWKPMIGGIGGFVKVGQDQILIEI	122
Qy	152	LNNKVRATIMTGDPINIFGSGNLTALGMSLNLPAKVEPIKMLKPGKDGKPLKROWPLT	211
Db	123	CGHKAIGTVLGGPVPVNIIGRNLLTQIGCTLNFPSPIETVPVKLPGMGDPKVKQWMP	182
Qy	212	KEKTEALKECEKMEQGLBEAPPTNPYNTPTFAIKKKDKNKRMLIDFRELNVQTDF	271
Db	183	EEEKALVEICTEMEKEGKISKIGPENPYNPVPAIKKDSKKRWKLVDFRELNKTQDF	242
Qy	272	TEIOLGPHPAGLAKKRITVLVDGDAVFSPLHEDRPYTAFTPLSVNNAEPGKRYIK	331
Db	243	WEVOLGPHPAGLAKKQKSVTVLVDGDAYSFVPLDKDFRYTAFTTIPSINNETPGIYQYN	302
Qy	332	VLPOGWKGSIPAIFQHTMROVLEPRKANKDVIIIOYMODILLASDORTLEHDRVVLQKE	391
Db	303	VLPOGWKGSIPAIFQCSMTFKLLEPRFKQPDVIYIYMODLYVGSOLETGOHRTKTEELHQ	362
Qy	392	LLNGLFGSTPDEKFOKDPYPYHWMGYELMPTKWLQIOLPQKEIWTAVNDIOKLVLGNWA	451
Db	363	HLRWGFTTDPKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTAVNDIOKLVLGNWA	422
Qy	452	AQLYPGKTKHLRLIRKMTLITEVQWTELAEELEBNRIILSQQBGGHYQOEBKELEA	511
Db	423	SOIYAGIKVRQCLLRGTCALTTEVPLTEEAEELEAENREILKEPHVGGYVDPKDLJIA	482
Qy	512	TVQKDOENOWYKIHQEBE-KTLKVGKYAKVKNTHNGIRLLAQVYOKTGEALVLTWRIP	570
Db	483	EIQOGOGQWYQIYIQEFPKLNKTKYARMGAHTNDVKQJTEAVOKIATESIVLWGTIP	542
Qy	571	KFHLPREIWEQWMDNWQVTPWDWFSVTPPLVRLAFNLVGDPPIPAETFFYDGCN	630
Db	543	KFKLPIQKETWEANWTYEQATWIPWEFVNTPLVLKWLQYOLEKEPELIGAEFFYVDGAAN	602
Qy	631	QSKREGAGYVTDGKVKVKLEQTTNOQALEAFAMALTDGSGPKVNIIVDSQYVWGISA	690
Db	603	RETKLGAGYVTDGRQGVVPLTDTTNQKTELQAIHLAQDSGLSEVNITWDSQYALGIQ	662
Qy	691	SOPTESKIVNOIIEEMIKKEAIVAVPAHKGIGGQNOVDHLVSOGIROLVFLTEKTEP	750
Db	663	AQDPKSESELVSQIIEILIKKEKYLAWPAHKGIGGQNOVDGLVSAGIRKVLFDGIDK	722
Qy	751	AQEEHEKYHSNVKLSHFGIPNLVAVINSQAQCOQKGEAIGHQVNAELGTWQMDCTH	810
Db	723	AQEEHEKYHSNWRAMASDFNLPVVAKEIVASCDCQLKGEAMHGQVDCSPGIMQLDCTH	782
Qy	811	LEGKIIIVAVHVASGFTAEAVIPOBSGRQATFLFLKLKLSRWPIITHLTDNGANFTSQEVK	870
Db	783	LEGKVIIVAVHVASGYIEAIVPAETGQETAYFLLLKLAGRPVKVTHTDNGSNFTSTTVK	842
Qy	871	MVANWIGIEQSGFYVPYQOSGVVEAMNHLKNOISRIREQANTTETVLMIAHCMNFKR	930
Db	843	AACWMAKIQEFGIPYNPQSGVIESMKEKLLKIIQGVDRDQAEHLKTAVQMAVFTIHFNR	902
Qy	931	RGGIGMDTPSSRLNMIITEQIEIOLQAKSNLKDFRYVIFREGRDQLWKGPCELLWKGE	990
Db	903	KGGIGGYASAGERIVDILATDITQRELOQITKIQNFRVYRDSRDPVWKGPAPKLLWKGE	962
Qy	991	AVLVKVGTDIKIIPRRKAKIITRDYGGROEMDSGSHLEGARD	1032
Db	963	AVTQDNDSIDKVPVRRKAKIITRDYQKQM---AGDDCVASRD	1001

RESULT 36
AAP60347
ID AAP60347 standard; protein; 1015 AA.
XX
AC AAP60347;
XX
DT 25-MAR-2003 (updated)

DT	01-JAN-1980	(first entry)	
XX	HTLV-III virus (HIV virus) pol protein.		
DE	XX		
XX	HTLV-III; HIV virus; AIDS; active immunization; pol protein;		
KW	passive immunization; vaccine; ss.		
KW	XX		
OS	HIV virus (HTLV-III).		
XX	XX		
PN	EP185444-A.		
XX	XX		
PD	25-JUN-1986.		
XX	XX		
PF	10-OCT-1985; 85EP-0307260.		
XX	XX		
PR	23-JAN-1985; 85US-0693866.		
PR	10-OCT-1984; 84US-0659339.		
XX	XX		
PA	(CENZ) CENTOCOR INC.		
PA	(USSH) NAT INST OD HEALTH.		
XX	XX		
PI	Chang NT;		
XX	XX		
DR	WPI; 1986-163443/26.		
DR	N-PSDB; AAN60240.		
XX	XX		
PT	New immunoreactive HTLV-III polypeptide expressed by transformed		
PT	cells - and derived antibodies, useful for diagnosis of AIDS and		
PT	in active or passive immunisation		
XX	XX		
PS	Disclosure; Fig. 3; 60pp; English.		
XX	XX		
CC	HIV virus cDNA is cleaved with restriction endonucleases to produce		
CC	the gag protein. The resulting protein and antibodies against it		
CC	are useful for immunoassay of HIV virus, e.g. by sandwich type RIA.		
CC	The protein may also be used in vaccines for active immunization.		
CC	See also AAP60346, AAP60348-9.		
CC	(Updated on 25-MAR-2003 to correct PA field.)		
XX	XX		
SQ	Sequence 1015 AA;		
	Query Match 57.1%; Score 3112; DB 7; Length 1015;		
	Best Local Similarity 56.2%; Pred. No. 4.8e-241;		
	Matches 563; Conservative 183; Mismatches 249; Indels 6; Gaps 31		
QY	33 SGSSSGSTGEIYAAREKTERAERTIOGSDRGLTAPRAGGDTIOCATNRGLAAPQFSLWK 92		
Db	18 SEQTRANSPTISSQTRANSPTREELQWGRDNNSPSEAGADROGTVSFNF--PQITLWQ 75		
QY	93 RPVVTAIEQOPVEVLDTGADDSIVAGIELGNNSPKIYGGIGGFINTKEYKNVEIYL 152		
Db	76 RPLVTIKIGQLKEALDGTADDTVLEMSLPGRWKPKMIGGIGGFKVROYDQILIEIC 135		
QY	153 NKKVRATIMGTDPINIFGRNLTALGMSLNLPAKVEPIKIMLPGKDGPKLRQWPLTK 212		
Db	136 GHKAIGTVIGPFPVNIIGNLLTQIOCTLNFPISPIETVPVKLPGMDGPKVQWPLTE 195		
QY	213 EKIEALKEICEKMEKEQLEAPPTNPYNTPTFAIKKKDKNKRMLIDFRELNKVTQDFT 272		
Db	196 EKIKALVEICTEMEKEGKISKIGPENPNTVPFAIKKKDKSTKRKLVDFRELNKRTQDEF 255		
QY	273 EIQLGIPHPAGLAKKRITVLVDGDAYFSIPLHEDFRPYTAFTLPSSVNNAPGKRYIKV 332		
Db	256 EVOLGIPHPAGLKKKSVTVLVDGDAYFSVPLDDEFKRYTAFTIPSNNETPGIRYQYNV 315		
QY	333 LPQGWKSPAIFOHTMRQVLEPFRKANCKDVIIQYMDDLIIASDRTDLEHDRVVLQKL 392		
Db	316 LPQGWKSPAIFOSSWTKILEPFPKQNPDIVIYQYMDLVGSDLEIGQHTKIEELRQH 375		
QY	393 NLGLGSTPDEKQKDDPYHMGVELMPTKWKLOKIQLPKEIKWTVDNDIQKLVGLNMAA 452		
Db	376 LLRWGLTTPDKKIQKEPFTLWMGVELHPDKWTQPIVLPEKDSVTVDNDIQKLVGLNMAA 435		

Db 631 LGCTLNFPISPIETVPVVKLPGMGPKVKQWPLTEEKIKALTAICDEMEKEGKITKIGPE 690
 QY 238 NPYNTPTFAIKKDKNKKWMLIDFRELANKVTQDTEIOIGLPHAGLAKKRITVLDVGD 297
 Db 691 NPYNTPTFAIKKDKSTKWKRLVDFRELANKRTQDFWEVOLGIPHPAGLAKKKSVTVDVGD 750
 QY 298 AYSFSLPHEDFRPYTAFTPLPVNNAEPGRKYIYKVLPGWGKSPAFQFMRQVLEPRFK 357
 Db 751 AYSFSLPYEDFRKYTAFTIPSRNNETPGIRYQYNNVLPQGWKSLAIFQSSMTKLEPRFK 810
 QY 358 ANKDVIIIOYMDLIIASDRDLEHDRVVLQKELLNGLGFSTPDEKFKQKPPYHWMGYE 417
 Db 811 QNPQIYQYMDLIIYVGSDELIGQHRTKIEELRQHLLRWGFTTPDKK-HKEPPFLWMGYE 869
 QY 418 LWPWKWKLOKQLPKOKETWVNDIQKLVGNLWAAQYPGKIKHLCLIRGLKMTLTREV 477
 Db 870 LHPDKWTVQTLPEKDSWTVNDQKLVGLKNWASQIYPGKIKVROLCKLLRGAKALTDIV 929
 QY 478 QWTELAELAELENRIILSQEGHYQBEKELEATVQKQDNQWYTKIHQEE-KILKYGK 536
 Db 930 PLTEAELELAENREILKEPVHGVYDPSKDLIAEIQOGQEQWYQIYQEPFKNLTKG 989
 QY 537 YAKVKNTHNGIRLLAQVQKIGKEALVINGRIPKFLHPVEREIQWQWMDNYQVWTPD 596
 Db 990 YAKWRTANDVKQLEATEAVQIAMEGIVINGKTPKFLPIQKETWETWTDYQWATWPE 1049
 QY 597 WDFVSTPLVRLAFNLVGDPIPGAETFTYDSCNRQSKGAGYVTDGKDKVKKLEQTT 656
 Db 1050 WEFVNTPLVRLWQLEKDPVGVETFYVDGAANRETKIGAGYVTDGRRKKIYSLTETT 1109
 QY 657 NQAELEAFAMALDTSQPKVNIIVDSQYVMGISAQPTSESKIVNQIIEEMIKKEATYV 716
 Db 1110 NQKTELOAICIALQDSGSEVINVDQYALGIIQAQPKDSESELVNIIEQLMKERVYL 1169
 QY 717 AWVPAHKGIGQNEVDHLVSGIRQVLFLEKIEPAQEEHEKYHNSVKLSHKFGIPNLVA 776
 Db 1170 SWVPAHKGIGNEQVDKLVSSGIRKVLFDGIDKAEHEHEKYHNSWRAMASDFNLPTVA 1229
 QY 777 RQVNSCAQCKQGEAIGHQVNAELGTWQMDCTHLEGGKIIIVAVHVASGFTEAEVPOES 836
 Db 1230 KEIVASDCOLKGEAMHGQVDCSPGIWQLDCTHLEGLIIVAVHVASGYWEAIVPAET 1289
 QY 837 GROTALFLKLASRPITHLTDNGANFTSQEVKMAVWIGIEQSFQYPPNPQSGVVEA 896
 Db 1290 GOETAYFTLKLRWPKVVIHTDNGSNFTSTAVAAACWAGIQOEFQIPYSPQSGVVEA 1349
 QY 897 MNHLKKNQISRIEQANTIEFIVLMAHCFMFKRGGIGDTPSERLINMTTQEQIQL 956
 Db 1350 MNKELKLIQVRDQAEHLKTAFLVMAVFIHFKRGGIGGYSGAGERIIDIIATDIQTREL 1409
 QY 957 QAKNSKLKDFRVYFREGDQLWKGPELLWKGAGVLVKVGTDTIKIIPRRKAKIIRDXG- 1015
 Db 1410 QKQITKQNFVRYVYRSDPSWKGPAKLLWKGEGAVIIDQNSDIKVPRRKAKIIRDXG 1469
 QY 1016 -----GREQMD 1021
 Db 1470 QMAGADCAGRQDED 1484

RESULT 40

AAO19387
 ID AAO19387 standard; Protein; 1003 AA.

XX AC AAO19387;

XX DT 10-DEC-2002 (first entry)

XX Lymphadenopathy-associated virus pol protein.

XX DE Lymphadenopathy-associated virus; LAV; HIV; human immunodeficiency virus;
 XX KW Lymphadenopathy; AIDS; gag; pol; orf Q; env; orf F; anti-HIV; virucide;
 XX KW vaccine.

Lymphadenopathy-associated virus.

US6440657-B1.

PD 27-AUG-2002.

XX 06-JAN-2000; 2000US-0478492.

XX 16-NOV-1984; 84GB-0029099.

PR 18-OCT-1985; 85CA-0493377.

PR 03-NOV-1992; 92US-0970954.

PR 17-OCT-1986; 86US-0920119.

PR 01-AUG-1989; 89US-0390499.

PR 06-DEC-1990; 90US-0622278.

PR 20-AUG-1991; 91US-0747506.

PR 01-OCT-1993; 93US-0130565.

PR 06-JUN-1995; 95US-0466920.

PR 05-DEC-1983; 83US-0558109.

PR 28-FEB-1985; 85US-0706562.

PR 30-AUG-1985; 85US-0771230.

PR 30-AUG-1985; 85US-0771247.

PR 30-AUG-1985; 85US-0771248.

XX (INSP) INST PASTEUR.

PA (CNRS) CENT NAT RECH SCI.

XX Montagnier L, Krust B, Chamaret S, Clavel F, Chermann J;

PI Barre-sinoussi F, Alizon M, Sonigo P, Cole S, Danos O;

PI Wain-Hobson S;

XX WPI; 2002-711525/77.

DR N-PSDB; AAL49920.

XX Novel chemically synthesized Human Immunodeficiency Virus envelope protein, useful in vitro diagnostic method for the detection of the presence or absence of antibodies that bind to antigens of HIV-1 retrovirus

XX Disclosure; Fig 3; 31pp; English.

XX The present invention relates to a chemically synthesised env peptide of Human immunodeficiency Virus (HIV) of less than 150 amino acid residues. Such a peptide is useful in an in vitro diagnostic method for the detection of the presence or absence of antibodies that bind to antigens of HIV-1 retrovirus, by contacting the peptide with a biological fluid for a time and under conditions sufficient for the peptide and antibody in the biological fluid to form a peptide-antibody complex, detecting the formation of the peptide-antibody complex by comparing the formation of peptide-antibody complex formation with a control sample, where the formation of peptide-antibody complex is correlated with the presence of antibodies that bind to antigens of HIV-1 retrovirus in the biological sample. It can also be used for preparation of vaccine compositions against AIDS or related syndromes. The present sequence is the Lymphadenopathy-associated virus LAV pol protein. The LAV virus was isolated from a patient with AIDS.

XX Sequence 1003 AA;

Query Match 57.1%; Score 3111; DB 23; Length 1003;

Best Local Similarity 56.6%; Pred. NO. 5.7e-241;

Matches 567; Conservative 176; Mismatches 235; Indels 24; Gaps 5;

QY 52 RAERETIQGSDRLTA--PRAGGDT-----IQGATNRGLAA-----PQFSIWL 91

Db 3 REDLAFLOGKAREFSSEQTRANSPTRRRELQVWGRDNNLSBAEADROQTVSFNFPQITLM 62

QY 92 KRPVVYATIEGQVPVEVLDTGADDSIVAGIEIGNNYSKPIVGIGGFTKEYKNVEIEV 151

Db 63 QRPLVTIKIGQLKEALLDTGADDTVLEEMSLPGRWKPKMIGGIGGFTKVRQYDILIEI 122

QY 152 LNNKVRATIMTGDPTINIFGRNLTALQMSLNLPAKVEPIKIMLPGKDGPKLRQWPLT 211

Db 123 CGHKAIGTVLVGPTPVNIIGNLLLTQIGCTLNFSPISPIETVPVVKLPGMGDPKVKQWPLT 182

Db 843 AACWAGIQEFGIPYNQSGVYVSMNKLKIIIGIIRDAEHLKTAQVMAVFIHFKR 902

Qy 931 RGIGDWTSERLINMITTEOEIQLQAKNSKLKDFRVYFREGDQLWGPGLLWKGEG 990

Db 903 KGGIGYSGAGERTIDIATDIQRELQKQIIRKIQNFRVYRDRSDPVWKGPAKLLWKGEG 962

Qy 991 AVLKVGKTDIKIIPRKAIRIDY-----GROEMD 1021

Db 963 AVVIQDNSEIKVPRKAKIIRIDYGKQAGMDCCVAGRQDED 1003

RESULT 44

ABU57552

ID ABU57552 standard; Protein; 1003 AA.

XX

AC ABU57552;

XX

DT 09-APR-2003 (first entry)

DE AIDS associated retrovirus-2 (9B) (ARV-2 (9B)) POL protein.

XX

KW Immunoassay; human immunodeficiency virus; HIV; antibody;

KW gag antigen; acquired immunodeficiency syndrome; AIDS; HIV infection;

KW vaccine; gene therapy; protein expression system; ARV-2;

KW AIDS associated retrovirus 2.

XX

OS Retroviridae.

XX

PN US6458527-B1.

XX

PD 01-OCT-2002.

XX

PF 28-JUN-1993; 93US-0083391.

XX

PR 24-DEC-1987; 87US-0138894.

PR 17-AUG-1992; 92US-0931191.

PR 31-OCT-1984; 84US-0667501.

PR 30-JAN-1985; 85US-0696534.

PR 06-SEP-1985; 85US-0773447.

XX

PA (CHIR) CHIRON CORP.

XX

PI Luciw PA, Dina D, Steimer K, Pescador RS, George-Nascimento C;

PI Parkes D, Halliwell R, Barr PJ, Truett M;

XX

DR WPI: 2003-182063/18.

DR N-PSDB; ABX77297.

XX

PT Immunoassay for detecting the presence of antibodies to a human

PT immunodeficiency virus (HIV) for diagnosing, treating or preventing HIV

PT infection, comprises contacting the sample with an HIV gag antigen -

XX

PS Disclosure; Fig 5B-E; 101pp; English.

XX

CC The invention describes an immunoassay for detecting the presence of

CC antibodies to a human immunodeficiency virus (HIV) in a sample by:

CC contacting the sample with an HIV gag antigen that binds anti-HIV

CC antibodies in an acquired immunodeficiency syndrome (AIDS) patient sera;

CC and determining binding, where the improvement comprises employing as

CC the gag antigen either a synthetic polypeptide or a recombinant

CC polypeptide. The method is used for detecting the presence of antibodies

CC to a human immunodeficiency virus (HIV) in a human sample. The

CC method is used in the preparation of a medicament for diagnosing,

CC treating or preventing HIV infection e.g. a vaccine or for gene

CC therapy. This is the amino acid sequence of a retroviral protein

CC isolated from AIDS associated retrovirus 2 (ARV-2) or HIV.

XX

SQ Sequence 1003 AA;

Query Match 57.0%; Score 3109; DB 24; Length 1003;

Best Local Similarity 56.8%; Pred. NO. 8.2e-241;

Matches 565; Conservative 179; Mismatches 243; Indels 12; Gaps 6;

RESULT 44

AAU84391

ID AAU84391 standard; Protein; 995 AA.

XX

Qy 38 GSTGEIYAAREKTERAERETIQ--GSDRGLTAPRAGGDTIQAGATNRGLAAPFSLWLRP 94

Db 11 GKARESSQTRANSPTRELRQVWGGENSL--EAGADR-QGTVSFNF--POITLWQRP 65

Qy 95 VVTAYIEGQVEVLLDTGADDSIVAGIELGNNSYSPKIVGGIGGINTKRYKNVEIEVLNK 154

Db 66 LVTIRIGQLKEALLDGTADDTVLEBMLPGKWKMGXGGFKVRQVDQIPVETCGH 125

Qy 155 KVRATMTGDTPIINIFGRNLTALGMSLNPVAKVEPIKMLPGKDGPKLRWPLTKK 214

Db 126 KAIGTVLVGPTPVNIIIGRNLITQIGCTLNPIETVTPVKLPGMDGPKVKOMPLTEK 185

Qy 215 IEALKEICEKMEKEGLEEAPPNTPTFAIKKDKNKRMLIDIFRELNVKVTQDFTEI 274

Db 186 IKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKDKSTKWRKLVDFRELNRKTQDFEW 245

Qy 275 QLGIPHPAGLAKRRITVLDVGDAYSIFSLHEDFRPYTAFTLPVSNNAEPKRYIYKVL 334

Db 246 QLGIPHPAGLAKRRITVLDVGDAYSIFSLHEDFRPYTAFTLPVSNNAEPKRYIYKVL 305

Qy 335 QGWKGSPTAFORTMROVLEPRKANKDVIIIOYMDLILASDRTDLEHDRVVLQKELLN 394

Db 306 QGWKGSPTAFORTMROVLEPRKANKDVIIIOYMDLILASDRTDLEHDRVVLQKELLN 365

Qy 395 GLGFSTPDEKFKOPPPYHMGYELWPTKWLQKIQLPQKEIWTVDIQLKLVGNMAAQL 454

Db 366 RRGFTTPDKKHOKEPPLWGYELHDPKWTVPIMLPEKDSMTVNDIQLKLVGNMAAQL 425

Qy 455 YPGIKTKHLRLIRKMTLTTEVOWTELEAELEENRIILSQEQEGHYQEEKELEATVQ 514

Db 426 YAGIKVKQLCKLIRGTALTEVTEAELEAELEENRIILSQEQEGHYQEEKELEATVQ 485

Qy 515 KQENOMTYKIHQEE--KILKVGKYAKVKNTHNGIRLLAQVYVOKICEALVINGRIPKPH 573

Db 486 KQGGQWYTIQYQEPFNKLTGKIRKRGHNDVKQLTEAVQGVSTESIVIMGKIPKFK 545

Qy 574 LPVEREINQWMDNYQWTVIPDWDFVSTPPLVRLAFNLVGDPIPGAETFTYDSCNRQS 633

Db 546 LPIQKETWEAWMEYQWATWIPDWDFVSTPPLVRLAFNLVGDPIPGAETFTYDSCNRQS 605

Qy 634 KEGKAGYVTRDGRKDKVKKLEQTTNOQAELEAFAMALTDSGPKVNIIVDSQVYNGISASQP 693

Db 606 KLGKAGYVTRDGRKDKVKKLEQTTNOQAELEAFAMALTDSGPKVNIIVDSQVYNGISASQP 665

Qy 694 TESEKIVNQIITEEMIKKEAIVAVVPAHKGIGGNOVDHLSQIRVQLFLEKIPPAQE 753

Db 666 DKSESELVQIIEQLIKKEKYLAVVPAHKGIGGNOVDHLSQIRVQLFLEKIPPAQE 725

Qy 754 EHEKYHSNVELSHKFGIPNLVARQIVNSCAQOQKEAIGHQVNAELGTWMDCTHLEG 813

Db 726 EHEKYHSNWRAMASDFNLPPVAKETIVASCDCQKLGEMHGVDCSPGIWQLDCTHLEG 785

Qy 814 KIIVAVHVASGFIEAEVPOESGROTALFLKLASRPVTHLHTDNGANFTSQEVKMA 873

Db 786 KIIVAVHVASGFIEAEVPOESGROTALFLKLASRPVTHLHTDNGANFTSQEVKMA 845

Qy 874 WTIIEQSGFVYPNPOSGVVEAMNHLKNQISRIREQANTIEIVLMATHCNFKRRGG 933

Db 846 WTIIEQSGFVYPNPOSGVVEAMNHLKNQISRIREQANTIEIVLMATHCNFKRRGG 905

Qy 934 IGDMPSERLINMITTEQEIQLQAKNSKLKDFRVYFREGDQLWGPGLLWKGAVL 993

Db 906 XGYSAGERIVDIATDIQTKELQKQIKIQNFRVYRDKDPLWKGPAKLLWKGAVV 965

Qy 994 KVGTDIKIIPRKAIRIDYGGQMDSGSHLEGARE 1032

Db 966 IQDNDSIKVPRKAKIIRIDYGGQMDSGSHLEGARE 1001

infection inhibitor; AIDS associated retrovirus; p24; gp41.

Human lymphotropic virus type III.

Key	Location/Qualifiers
100	100
101	101
102	102
103	103
104	104
105	105
106	106
107	107
108	108
109	109
110	110
111	111
112	112
113	113
114	114
115	115
116	116
117	117
118	118
119	119
120	120
121	121
122	122
123	123
124	124
125	125
126	126
127	127
128	128
129	129
130	130
131	131
132	132
133	133
134	134
135	135
136	136
137	137
138	138
139	139
140	140
141	141
142	142
143	143
144	144
145	145
146	146
147	147
148	148
149	149
150	150
151	151
152	152
153	153
154	154
155	155
156	156
157	157
158	158
159	159
160	160
161	161
162	162
163	163
164	164
165	165
166	166
167	167
168	168
169	169
170	170
171	171
172	172
173	173
174	174
175	175
176	176
177	177
178	178
179	179
180	180
181	181
182	182
183	183
184	184
185	185
186	186
187	187
188	188
189	189
190	190
191	191
192	192
193	193
194	194
195	195
196	196
197	197
198	198
199	199

Misc-difference 50
/label= Leu, Pro

```
Misc-difference 615
/label= Lvs. Arg.
```

Misc-difference 956

US5853978-A.

29-DEC-1998.

29-JUL-1994: 94US-0282857.

08-MAY-1986 86US-0861016

24-DEC-1984; 84US-0685272.

02-AUG-1988; 88US-0227568.

19-NOV-1992; 92US-0979391;
29-SEP-1993; 93US-0129009;

29-JUL-1994; 9405-0282857.

(GETH) GENENTECH INC.

Berman PW, Capon DJ, Lasky

WPI; 1999-094894/08.

N- F3DB; HAV/42/1.

New AIDS-associated
of AIDS or to inhibi

dissemination of

Example 1; Fig 2A-J; 47pp; English.

This invention describes a novel fusion polypeptide which comprises

polypeptide, that specifically binds complementary antibody and (b)

polypeptide. Also described are (1) a fusion polypeptide having

herpes simplex virus (HSV) gp signal peptide sequence fused in reading frame with a polypeptide sequence other than HSV gp, (2) a nucleic acid encoding a fusion polypeptide as in (1); (3) an expression vector comprising a nucleic acid as in (2); (4) a host comprising a vector as in (3). The AIDS-associated fusion polypeptides can be used in the detection of and vaccination against viral etiological agents of AIDS. They can also be administered as a pharmaceutical agent to inhibit infection by AIDS associated retrovirus or dissemination of such retrovirus in infected individuals.

XX	Sequence	1012 AA:
SQ	Sequence	

Query Match 56.9% Score 3104.5 DB 20: Length 1012:

Best Local Similarity 56.0%; Pred. No. 1.9e-240;
Matches 567. Conservative 179. Mismatches 250.

[illegible][illegible]

DD 1 GKSAFLQQAKREFSSEQIKANSFIISSEQIKANSFIKRELQVWGRDNNSSASEAGADKQGI 60

QY 84 AA--PQFSLWKRPPVVTAYIEGQPVEVLDTGADDSIVAGIELGNNSYPKRVGGIGFIN 140

Db 61 VSFNFPQITLWQRPLVTIKIGGQLKEALLDTGADDTVLEEMSLPGRWPKMIGGIGGFIK 120

QY 141 TKEYKNVEIEVLNKKVRATIMTGDTPINIFGRNILTALGMSLNPVAKVEPIKIMLKPGK 200

Db 121 VRQYDQILIEICGHAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLKPGM 180

QY ,201 DGPKLQWPLTKKEIEALKEICEKEKEGQLEEA PPNPYNTPTFAIKKKDKNKRMLID 260

[illegible]

Search completed: September 25, 2003. 12:02:16

Job time : 64 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2003, 12:01:21; Search time 21 Seconds
(without alignments)
2087.336 Million cell updates/sec

Title: POL_HV2RO
Perfect score: 5452
Sequence: 1 TGRFFRTGLGKEAPQLPRG.....RQMDSGSHLEGAREDGEMA 1036

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 99%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTOUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5075.5	93.1	3080	6	5223423-4
2	5057.5	92.8	1055	2	US-08-659-251-5
3	5057.5	92.8	1055	3	US-09-256-490-5
4	5057.5	92.8	1055	5	PCT-US96-11445-5
5	3324	61.0	1018	4	US-09-206-551-46
6	3301	60.5	1016	4	US-09-625-972-24
7	3155	57.9	1003	2	US-07-743-357-9
8	3150.5	57.8	1014	4	US-09-319-588C-6
9	3119	57.2	1004	2	US-07-743-357-7
10	3117	57.2	1005	2	US-07-743-357-1
11	3116	57.2	1016	2	US-07-743-357-3
12	3112	57.1	1003	4	US-09-309-572-17
13	3112	57.1	1015	3	US-08-463-210-9
14	3112	57.1	1015	3	US-09-124-900-3
15	3112	57.1	1016	2	US-07-743-357-2
16	3109	57.0	1016	2	US-07-743-357-5
17	3107	57.0	1016	2	US-07-743-357-4
18	3105.5	57.0	1003	2	US-07-743-357-8
19	3095.5	56.8	1003	2	US-07-743-357-10
20	2999	55.0	913	2	US-07-743-357-22
21	2957	54.2	913	2	US-07-743-357-6
22	1921.5	35.2	562	3	US-09-117-217-14
23	1921.5	35.2	562	4	US-09-735-487-14
24	1889.5	34.7	560	4	US-09-752-652-1
25	1810	33.2	1150	3	US-09-238-303-9
26	1810	33.2	1150	4	US-09-946-239-9
27	1673	30.7	1031	4	US-08-811-682-15

28	1455.5	26.7	427	4	US-09-690-265-1	Sequence 1, Appli
29	1068	19.6	314	1	US-08-589-446-6	Sequence 6, Appli
30	1068	19.6	314	1	US-08-444-882-6	Sequence 6, Appli
31	1068	19.6	314	1	US-08-389-459A-6	Sequence 6, Appli
32	1068	19.6	314	3	US-08-987-867A-6	Sequence 6, Appli
33	1055	19.4	327	4	US-08-679-493A-72	Sequence 72, Appli
34	1048	19.2	327	4	US-08-679-493A-71	Sequence 71, Appli
35	948	17.4	261	6	5320958-3	Patent No. 5320958
36	924.5	17.0	290	4	US-09-231-182B-2	Sequence 2, Appli
37	828.5	15.2	770	1	US-07-648-796A-8	Sequence 8, Appli
38	828.5	15.2	799	1	US-07-648-796A-7	Sequence 7, Appli
39	827.5	15.2	376	1	US-07-648-796A-6	Sequence 6, Appli
40	821	15.1	275	6	5252477-3	Patent No. 5252477
41	804	14.7	247	1	US-07-648-796A-73	Sequence 2, Appli
42	691	12.7	237	4	US-08-679-493A-73	Sequence 73, Appli
43	586.5	10.8	917	4	US-08-259-451-11	Sequence 11, Appli
44	576.5	10.6	1203	3	US-09-075-272-4	Sequence 4, Appli
45	554.5	10.2	1737	4	US-09-309-572-13	Sequence 13, Appli

ALIGNMENTS

RESULT 1
5223423-4
; Patent No. 5223423
; APPLICANT: FRANCHINI, GENOVEFFA; WONG-STAAAL, FLOSSIE;
; GALLO, ROBERT
; TITLE OF INVENTION: CHARACTERIZATION OF REPLICATION COMPETENT
; HUMAN IMMUNODEFICIENCY TYPE 2 PROVIRAL CLONE HIV-2 SBL/ISY
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/331,212
; FILING DATE: 03-31-1989
; SEQ ID NO: 4:
; LENGTH: 3080
5223423-4

Query Match	93.1%	Score	5075.5	DB	6	Length	3080
Best Local Similarity	91.7%	Pred. No.	0				
Matches	950	Conservative	49	Mismatches	36	Indels	1
Gaps	1						
QY	1	TGRFFRTGLGKEAPQLPRGSSAGADTNSTPSGSSSGTGEIYAAREKTERAERTIQG	60				
Db	584	TGWFFRANTMCKEAPQLPRGPKFACANTINSPNGSSGPTGEVHAAREKTERAETKIQR	643				
QY	61	SDRGLTAPRAGDITTOGATNRGLAAPQSLWKRPVVTAYIEGQPVVLLDTGADDSIVAG	120				
Db	644	SDRGLAASRRARDTQ-RDDRGLAAPQSLWKRPVVTAYIEDQPVVLLDTGADDSIVAG	702				
QY	121	IELGNYSKIVGGIGGFINTKEYKNVEVLNKKVRATIMTGDPTINFGRNLTALGM	180				
Db	703	IELGSNYSKIVGGIGGFINTKEYKDVLRVNLKKVRATIMTGDPTINFGRNLTALGM	762				
QY	181	SLNLPVAKVEPIKIMLKFCKGPKLRQWPLTKEKTEALKECEKMEKEGLEEAPPTNPY	240				
Db	763	SLNLPVAKIEPVKVTLPKCKGPKQKQWPLTREKTEALREICEKMEREGQLEAPPTNPY	822				
QY	241	NTPPTFAIKKDKNKKRMLIDFRELNVQDFTFIEQIGLPHAGLAKKRITVLDVGDYF	300				
Db	823	NTPPTFAIKKDKNKKRMLIDFRELNVQDFTFIEQIGLPHAGLAKKRITVLDVGDYF	882				
QY	301	SIPLHEDFRPYTAFTLPSVNNAPCKRYIKVLPQGWKSPAFIOTHRVLEPPFRKANK	360				
Db	883	SIPLHEDFRQYTAFTLPSVNNAPCKRYIKVLPQGWKSPAFIOTHRVLEPPFRKANKP	942				
QY	361	DVIIITQYDDILIASDRDLDLHDRVVLQIKELLNGLGFSTPDEKFKQPPYHWMGYELWP	420				
Db	943	DVIIITQYDDILIASDRDLDLHDRVVLQIKELLNGLGFSTPDEKFKQPPYHWMGYELWP	1002				
QY	421	TKWKLOKLTOLPOKEITWTVNDIQKLVGLNNAQYPGIKTKHLCKRLIRGKMTLTBEVQMT	480				
Db	1003	TKWKLOKLTOLPOKEITWTVNDIQKLVGLNNAQYPGIKTKHLCKRLIRGKMTLTBEVQMT	1062				

```
Qy 481 ELAAEAELEENRIILSQBQEGHYQEEKELEATVQKQENQWTKYKHQEEKILKVGKYAKV 540
Db 1063 ELAAEAELEENRIILSQBQEGHYQEEKELEATVQKQDNQWTKYKHQEEKILKVGKYAKI 1122
Qy 541 KNTHTNGIRLLAQQVQKIGKEALVWGRIPKFLPVERETWEQWNDYQWVTWIPDWFV 600
Db 1123 KNTHTNGVRLLAQQVQKIGKEALVWGRIPKFLPVERETWEQWNDYQWVTWIPDWFV 1182
Qy 601 STPLPLVRLAFNLVGDPIPGAETFTVDSGSCNRQSKGKAGVYVTDGRKDKVKKLEQTTNQQA 660
Db 1183 STPLPLVRLAFNLVGDPIPGAETFTVDSGSCNRQSKGKAGVYVTDGRKDKVKRILQTTNQQA 1242
Qy 661 ELEAFAMALTDGSPKVNIIVDSQVVMGISQSTESKIVNQIIEEMIKKEAIYVAVWP 720
Db 1243 ELEAFAMAVTDSGPKVNIVDSQVVMGIVTGPAAESERIVNKIIEEMIKKEAIYVAVWP 1302
Qy 721 AHKGIGNQEVHLSVQIGIRQVLFLEKIEPAQEHEKYSNVKELSHKFGIPNLVARQIV 780
Db 1303 AHKGIGNQEVHLSVQIGIRQVLFLEKIEPAQEHEKYSNVKELSHKFGIPNLVARQIV 1362
Qy 781 NSCAQCOOKEAETHGVNAELGTWQMDCTHLEKLIIVAVHVASGFTAEAVIPOESGRQT 840
Db 1363 NTCAQCOOKEAETHGVNAELGTWQMDCTHLEKLIIVAVHVASGFTAEAVIPOESGRQT 1422
Qy 841 ALFLLKLASRPITHLHTDNGANFTSOEVKVAWVWIGIEQSGFVYPNPQSQGVVEAMNHH 900
Db 1423 ALFLLKLASRPITHLHTDNGANFTSOEVKVAWVWIGIEQSGFVYPNPQSQGVVEAMNHH 1482
Qy 901 LKNOISRIRQANTETIVLMAHGMNFKRRGIGDMTPSERLINMITTEQETQFTQAKN 960
Db 1483 LKNOISRIRQANTETIVLMAHGMNFKRRGIGDMTPSERLINMITTEQETQFTQAKN 1542
Qy 961 SKLKDPRVYPRGRDQLWQPGELLWKGEGAVLVKVGTDKIIIPRRKAKIIRDYGROEM 1020
Db 1543 SKLKNFYPRGRNQLWQPGELLWKGEGAVLVKVGTDKIIIPRRKAKIIRDYGROEM 1602
Qy 1021 DSGSHLEGAREDEGEMA 1036
Db 1603 DSGSHLEGAREDEGEMA 1618

RESULT 2
US-08-659-251-5
; Sequence 5, Application US/08659251
; Patent No. 5883081
; GENERAL INFORMATION:
; APPLICANT: Kraus, Guenter
; APPLICANT: Wong-Staal, Flossie
; APPLICANT: Talbott, Randy
; APPLICANT: Poeschla, Eric
; TITLE OF INVENTION: Isolation of No. 5883081el HIV-2 Proviruses
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,251
; FILING DATE: No. 5883081 yet assigned
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/001,441
; FILING DATE: 26-JUL-1995
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Garrett-Wackowski, Eugenia
; REGISTRATION NUMBER: 37,330
; REFERENCE/DOCKET NUMBER: 02307E-056410US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1055 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..1055
; OTHER INFORMATION: /note= "pol protein encoded by HIV-2KR"
; US-08-659-251-5

Query Match 92.8%; Score 5057.5; DB 2; Length 1055;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 949; Conservative 43; Mismatches 43; Indels 1; Gaps 1;

Qy 1 TGRFFRTGLGKEAPQLPRGPPSAGADTNTSPSGSSSGSTGEIYAAEKTERAEREPTIQG 60
Db 21 TGNFFRDWPMGKEASQLPRDPSPAGADTNTSPSRSPAREVLAAREAEAEENETIQG 80
Qy 61 SDRLGLAPRAGGTIOGATNRGLAAPQFSLWKRPVVTAYTEGQPVVELLDTGADDSIVAG 120
Db 81 GDRGLTAPRTTRDTORG-DRGFAAPQFSLWKRPVVTAYVEGQPVVELLDTGADDSIVAG 139
Qy 121 IELGNYSKPIVGGIGGIFINTKEYKNVEIEVLNKKVRATIMTGDTPINIFGRNILTALGM 180
Db 140 IELGSNYSKPIVGGIGGIFINTKEYKNVEIKVLNKKVKATIMTGDTPINIFGRNILTALGM 199
Qy 181 SLNLPVAKVEPIKIMLKPGKDGPKLRQWPLTKKEIALKEICEKMEKEGQLEAPPTNPY 240
Db 200 SLNLPVAKVDPIKVLKPGKDGPKVRQWPLTKKEIALKEICEKMEREGQLEAPPTNPY 259
Qy 241 NPTFFAIKKKDKNKRMLIDFRELNVKVTQDFTFIQLGIPHPAGLAKRRITVLDVGDYF 300
Db 260 NPTFFAIKKKDKNKRMLIDFRELNVKVTQDFTFIQLGIPHPAGLAKRRITVLDVGDYF 319
Qy 301 SIPLHEDFRPYTAFTLPSVNNAPGKRYIKVLPQGMKGSPAIHQHTMRQVLEPFRKAN 360
Db 320 SIPLHEDFRQYTAFTLPTVNNAPGKRYIKVLPQGMKGSPAIHQHTMRQVLEPFRKAN 379
Qy 361 DVIIIOYMDILLIADSDRTDLEHDRVVLQKLKELLNGLGFSPTDPEKFKDPPYHMGYELWP 420
Db 380 DVILOYMDILLIADSDRTDLEHDRVVLQKLKELLNGLGFSPTDPEKFKDPPYHMGYELWP 439
Qy 421 TKWKLOKIQLPQKEIWTVDIOKLVLGNWAAOLYPGIKTKHLCLIRIGKMTLTVEQWT 480
Db 440 TKWKLOKIQLPQKEIWTVDIOKLVLGNWAAOLYPGIKTKHLCLIRIGKMTLTVEQWT 499
Qy 481 ELAAEAELEENRIILSQBQEGHYQEEKELEATVQKQDNQWTKYKHQEEKILKVGKYAKV 540
Db 500 ELAAEAELEENRIILSQBQEGHYQEEKELEATVQKQDNQWTKYKHQEEKILKVGKYAKI 559
Qy 541 KNTHTNGIRLLAQQVQKIGKEALVWGRIPKFLPVERETWEQWNDYQWVTWIPDWFV 600
Db 560 KNTHTNGVRLLAQQVQKIGKEALVWGRIPKFLPVERETWEQWNDYQWVTWIPDWFV 619
Qy 601 STPLPLVRLAFNLVGDPIPGAETFTVDSGSCNRQSKGKAGVYVTDGRKDKVKKLEQTTNQQA 660
Db 620 STPLPLVRLAFNLVGDPIPGAETFTVDSGSCNRQSKGKAGVYVTDGRKDKVKRILQTTNQQA 679
Qy 661 ELEAFAMALTDGSPKVNIIVDSQVVMGISQSTESKIVNQIIEEMIKKEAIYVAVWP 720
Db 680 ELEAFAMALTDGSPKVNIIVDSQVVMGIVAGQPTSESKLVNQIIEEMIKKETLYVAVWP 739
Qy 721 AHKGIGNQEVHLSVQIGIRQVLFLEKIEPAQEHEKYSNVKELSHKFGIPNLVARQIV 780
Db 740 AHKGIGNQEVHLSVQIGIRQVLFLEKIEPAQEHEKYSNVKELSHKFGIPNLVARQIV 799
```

Db 740 AHKGIGNQVDHLVSQGIQVLFLEKTEPAQEHEEKYHSNVKLSHFKPLVARQIV 799
QY 781 NSCAOQKGEAIGHOVNAELGTWQMDCTHLEGGKIIIVAVHVASFIEAEVIPSQSGROT 840
Db 800 NTCACQKGEAIGHQVDAELGTWQMDCTHLEGGKIIIVAVHVASFIEAEVIPSQSGROT 859
QY 841 ALFLKLSRWPITHLHTDNGANFTSQEVKVMVAMWIGIEQSGFVYPNPSQSGVVEAMNHH 900
Db 860 ALFLKLSRWPITHLHTDNGANFTSQEVKVMVAMWIGIEQSGFVYPNPSQSGVVEAMNHH 919
QY 901 LKNQISRIREQANTTETIVLMAHCHMNFRRGGIGDMTPSERLINMITTEQIEQIQLAKN 960
Db 920 LKNQISRIREQANTTETIVLMAHCHMNFRRGGIGDMTPSERLINMITTEQIEQIQLAKN 979
QY 961 SKLDFRYVFRGRDQWKPGELLWKGEGAVLVKVGTDIKTIIPRRKAKIIRDYGGROEM 1020
Db 980 SKLKNFRYVFRGRDQWKPGELLWKGEGAVLVKVGTDIKTIIPRRKAKIIRDYGGRRREV 1039
QY 1021 DSGSHLEGAREDEGEMA 1036
Db 1040 DSSSHLEGREDEGEVA 1055

RESULT 3

US-09-256-490-5
; Sequence 5, Application US/09256490
; Patent No. 6235881
; GENERAL INFORMATION:
; APPLICANT: Kraus, Guenter
; APPLICANT: Wong-Staal, Flossie
; APPLICANT: Talbott, Randy
; APPLICANT: Poeschla, Eric
; TITLE OF INVENTION: Isolation of No. 6235881el HIV-2 Proviruses
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/256,490
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/659,251
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett-Wackowski, Eugenia
; REGISTRATION NUMBER: 37,330
; REFERENCE/DOCKET NUMBER: 02307E-056410US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1055 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..1055
; OTHER INFORMATION: /note= "pol protein encoded by HIV-2KR"
US-09-256-490-5

Query Match 92.8%; Score 5057.5; DB 3; Length 1055;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 949; Conservative 43; Mismatches 43; Indels 1; Gaps 1;
QY 1 TGRFFRTGPLGKEAPQLPRGPSSAGADTNSTSGSSSGTSGEIIYAAREKTERAERTIQQ 60
Db 21 TGWFFRDWPMGKEASQLPDRPSPAGADTNSTSRSPAREVLAAREAEARENETIQQ 80
QY 61 SDRGLTAPRAGDGTTCAGATNRCGLAAPQSLWLRPVVTAYIEQOPVEVLLDTGADDSIYAG 120
Db 81 DGRGLTAPTRRDTTTQRG--DRGFAAPQFSLWLRPVVTAYIEQOPVEVLLDTGADDSIYAG 139
QY 121 IELGNYSPIKIVGGIGGFINTKEYKNVEIEVLNKKVRATIMTGDPIINFGRNLTALCM 180
Db 140 IELGNYSPIKIVGGIGGFINTKEYKNVELKVLNKKVKATIMTGDPIINFGRNLTALCM 199
QY 181 SUNLPVAKVEPIKMLKPGKDGPKLRQWPLTKETKEALKEICEKMEKEQLEBEAEPPTNPY 240
Db 200 SUNLPVAKVDPIKVLKPGKDGPKVRQWPLTKETKEALKEICEKMEREGQLEAEPPTNPY 259
QY 241 NPTTFAIKKKDKKWRMLIDFRELANKVTODFTETIOLGPHPAGLAKKRITVLDVGDYVF 300
Db 260 NPTTFAIKKKDKKWRMLIDFRELANKVTQEFTEIOLGPHPAGLAKKRITVLDIGDAYF 319
QY 301 SIPLHEDFRPYTAFTLPVNNAPCGKRYIKVLPQGWKGSIPAIFQHTMRQVLEPFRKAN 360
Db 320 SIPLHEDFRPYTAFTLPVNNAPCGKRYIKVLPQGWKGSIPAIFQHTMRQVLEPFRKAN 379
QY 361 DVIITQYMDDIILIASDRTDLEHDRVVLQKELLNGLGFSTPDEKFOKDPYPYHMGYELWP 420
Db 380 DVIIVQYMDDIILIASDRTDLEHDRVVLQKELLNGLGFSTPDEKFOKDPYPYHMGYELWP 439
QY 421 TKWKLOKIQLOPKEIWTNDIQKLVGLNWAALYPGIKTKHLCLRIEGLKMTLTEVQWT 480
Db 440 TKWKLOKIQLOPKEIWTNDIQKLVGLNWAALYPGIKTKHLCLRIEGLKMTLTEVQWT 499
QY 481 ELAAEAELENRIILSQEQEGHYIYQBEKELEATVQKQDNQWTKYKHQEKILKVKYAKV 540
Db 500 ELAAEAELENRIILSQEQEGHYIYQBEKELEATVQKQDNQWTKYKHQEKILKVKYAKI 559
QY 541 KNTHTNGIRLLAQQVQKIGKEALVITWGRIPKPHLPVEREINWQWMDNTHQVWIWDWDFV 600
Db 560 KNTHTNGIRLLAQQVQKIGKEALVITWGRIPKPHLPVEREINWQWMDNTHQVWIWDWDFV 619
QY 601 STPLVLRLAFNLVGDPIPCAETFTYDGCNROSKEGKAGYITDRGRDKVRILEQTTNQQA 660
Db 620 STPLVLRLAFNLVGDPIPCAETFTYDGCNROSKEGKAGYITDRGRDKVRILEQTTNQQA 679
QY 661 ELEAFAMALTDSPKVNIIIVDSQYVYMGISASOPTSESKIVNQIIEEMIKKEAIVAWVP 720
Db 680 ELEAFAMALTDSPKVNIIIVDSQYVYMGISASOPTSESKIVNQIIEEMIKKEAIVAWVP 739
QY 721 AHKGIGNQVDHLVSQGIQVLFLEKTEPAQEHEEKYHSNVKLSHFKPLVARQIV 780
Db 740 AHKGIGNQVDHLVSQGIQVLFLEKTEPAQEHEEKYHSNVKLSHFKPLVARQIV 799
QY 781 NSCAOQKGEAIGHOVNAELGTWQMDCTHLEGGKIIIVAVHVASFIEAEVIPSQSGROT 840
Db 800 NTCACQKGEAIGHQVDAELGTWQMDCTHLEGGKIIIVAVHVASFIEAEVIPSQSGROT 859
QY 841 ALFLKLSRWPITHLHTDNGANFTSQEVKVMVAMWIGIEQSGFVYPNPSQSGVVEAMNHH 900
Db 860 ALFLKLSRWPITHLHTDNGANFTSQEVKVMVAMWIGIEQSGFVYPNPSQSGVVEAMNHH 919
QY 901 LKNQISRIREQANTTETIVLMAHCHMNFRRGGIGDMTPSERLINMITTEQIEQIQLAKN 960
Db 920 LKNQISRIREQANTTETIVLMAHCHMNFRRGGIGDMTPSERLINMITTEQIEQIQLAKN 979
QY 961 SKLDFRYVFRGRDQWKPGELLWKGEGAVLVKVGTDIKTIIPRRKAKIIRDYGGROEM 1020
Db 980 SKLKNFRYVFRGRDQWKPGELLWKGEGAVLVKVGTDIKTIIPRRKAKIIRDYGGRRREV 1039
QY 1021 DSGSHLEGAREDEGEMA 1036


```

; ORGANISM: Simian immunodeficiency virus
; FEATURE:
; OTHER INFORMATION: Deduced amino acid sequence of SIVrcm Pol protein
us-09-206-551-46

Query Match      61.0%: Score 3324; DB 4: Length 1018;
Best Local Similarity 59.5%: Pred. No. 1.8e-275;
Matches 627; Conservative 156; Mismatches 209; Indels 62; Gaps 8;

QY  4  FFTGPGUGK-EAPQL-----PRGPSSAGADNTSPGSSSGST 40
DB  1  FFRETPLGSGEAEELSDFDLSDLRDGEQLRPCRDAKELSELSEGRGTKETTE----- 53
QY  41  GEIYAAREKTERAERETIQGSDRGLTAPRAGGDTIQGATNGLAAPFOSLWKRPVVTAYI 100
DB  54  ----AGREQEER-----GSIQSLNLPFALMKRPTTIAQI 85
QY  101 EGQPEVLLDTGADDSIVAGIELGNYSPTKIVGGIGGFINTKEVKNVEIEVLNKKVRAI 160
DB  86  EGQKVEVLLDTGADDTVIEGIELGNWPTKLIIGIGGYINVKYKNEIEIAGRTAHV 145
QY  161 MTGDPTINIEGRNILTALGMSLNPVAKVEPIKIMLPKGDGPKLRQWPLTKKEIEALKE 220
DB  146 LVGPTPVNIIGRNVLKIGATLNFPIISQETIKVELASGQDGPVKQWPLSKKEIALTE 205
QY  221 ICEKMEKEGQLEAPPTNPNTPTFAIKKKDKNKKRMLIDFRELUNKYVQDFTETQLGIPH 280
DB  206 ICNAMEKEGKISKIGPENPTNTPTFCIKKKDSTKWRKLVDPRFLUNKRTQDFFEVQLGIPH 265
QY  281 PAGLAKRRRTVLLDVGDAVFSIPLHEDFRPYTAFTLPSVNAEPGKYIVKYLPGQWKGS 340
DB  266 PGLGKQCERTVLLDGDAYSFCLLYEPFRKYTAFTIPAVANQGGPVRIQYNVLPGQWKGS 325
QY  341 PAIFQHTMRQVLPEPFRKANKDVIIQYMDDLTASDRTLDEHDRVVLQKLKELLGLGFST 400
DB  326 PAIFQASANKILQFPREENPDVVIYQYMDLFLVCGSDRTKLEHDKWIKQLRDHLLFWGFET 385
QY  401 PDEKFKQDPYPYHWGVELWPTKKLQKIQLPQKEIWTVNDIQKLVGVNLWAAQLYPIGKT 460
DB  386 PDKKFKQDPYPLMWGVELHPKSWTVQBIKLPKEKEEWTVNDIQKLVGKLNWASQIYSGIRT 445
QY  461 KHLCLIRIGKWTILTEEVQWTELAEELEENKRIILSOEGHYEQEKELEATVQKDOENQ 520
DB  446 KELCLKLRGAKALDEKVEMTKEAEIEYEENKMILKEKLGHYVYDEKPLVANIQKLGEGQ 505
QY  521 WTYKIIHQEE-KILKVGKIYAKVNHTHTNGIRLLAQVQKIGKEALVWGRIPKPHLPVERE 579
DB  506 WSYQIEQESGKPLKTGKYAKOKTAHTNEIRMLAGLVQKIAKEAIVWGRLLPTPLPIERE 565
QY  580 IWEQWMDNWTWQVTLIPDWDFVSPPLVRLAFNLVGDPIPGAENFYFTDGCNCRQSKEGKAG 639
DB  566 VWD-WRSQYVQVWTIPDWEEVSTPPLIRLGNLVKDPPIGPEEYVYVDGAANRNSKIKGAG 624
QY  640 YVTDRGKDKVKKLEQTNNOQAELEAFAMALTDSGPKVNIIVDSQYVNGVIGSASQPTESK 699
DB  625 YVTVNRGKVKYKLEETTNOQAELEAVLLALKDSGPKVNIIVDSQYVIGILEAQDPTSDSG 684
QY  700 IVNQIIEEMTKKEAIVYAWVPAHKGIGGNOQEVDLHVSQGIQRQVLFLEKIEPAQEHEKHYH 759
DB  685 LVTEIINQMIGKEAVLVSWVPAAHKGIGGNEVDKLVSKGIRQVLFLDGIEKAQEHEKHYH 744
QY  760 SNVKELSHKFGIPNLVARIQVNSCAQOQKEAGIAGOVNAELGTWQMDCTHLEGGKIIVA 819
DB  745 NNWRALAEDFQIPIQVAKETIVACQPKQVKEAIGHGVDAISPGTQWQMDCTHLEGGKIIVA 804
QY  820 VHVASGFIIEAEVIPSQSGROTALFLKLKASRWPTTLHTLTDNGANFTSOEYKVMVAWWTIGIE 879
DB  805 VHVASGYIEAEVIPAETGKETAHELLKLAAWPNVKRLHTDNGANFTSAAQVACVWAAQIE 864
QY  880 QSGFVYPNPQSQGVVEAMNHHLLKNQISRIREQANTTETIVLMAIHCNMFKRRGGIGDMTP 939
DB  865 HAFGVYNPQSQGVVEAMNKKQLKRIIEQVREQAEKLETAQVMAVLVFNFRKKGIGGYSA 924

```

```
QY 542 NHTINGIRLLAQVQKIGKEALVIWGRPKFHLPLVEREWEQWMDNYQVWTWIPDMDFVS 601
DB 527 NAHTNELRTLAGLVQKIAKECIVLWGLPKFYPLEREVWDQWHDYQVWTWIPWEFIS 586
QY 602 TPPLVRLAFNLVGDPIPGATFTYDSCNRSQSEKGRAGYVTDGRKDKVKLEOTNQAAE 661
DB 587 TPPLIRLWYNLLKEPIGEDVYVVDGAANRSKGRAGYVTDGRKDKVKLEOTNQAAE 646
QY 662 LEAFAMALTDGPKVNIIVDSQVVMGISASOPTSESKIVNOIIEEMIKKEALYVAVWPA 721
DB 647 LKALIELAKDSGRVNIIVTDSQALGILTASPOQSONPIVREINLMIAKEAYLSWVPA 706
QY 722 HKGIGNQEDHLVDSQIGQVLFLEKIEPAQEHEHYHSNKSLSHKGIPNLVARQIVN 781
DB 707 HKGIGNQEDHLVDSQIGQVLFLEKIEPAQEHEHYHSNKSLSHKGIPNLVARQIVN 766
QY 782 SCACQCKGGAIGHOVNAELGTWQMDCTHLEGLKIIIVAVHIVASGFTEAEVIPSQRQTA 841
DB 767 QCPKQIKGPPIHGOVDASPTWQMDCTHLEGLKIIIVAVHIVASGFTEAEVIPSQRQTA 826
QY 842 LFLKLKASRPITHLHTDNCANTSOEVKVMWVWIGIEQSGFYVPYNPQSGVVEAMNHL 901
DB 827 HFLKLKASRPITHLHTDNCANTSOEVKVMWVWIGIEQSGFYVPYNPQSGVVEAMNHL 886
QY 902 KNOISRIREQANTTETIVLMAIHCNMFRRGGIGDMTPSERLINMITTEQEIQLQAKNS 961
DB 887 KTIIEQVRDQAEKLETAVQMAVLIHNFRRKGGIGGYSPGERIVDIITDILTTLKQONIS 946
QY 962 KLKDFRYFREGDQWLKGGELLWKGEGAVLVKGTDKIIPRRKAKIIRDYGGREQMD 1021
DB 947 KIONFRYYREGDQWLKGGELLWKGEGAVLVKGTDKIIPRRKAKIIRDYGGREQMD 1004
QY 1022 SGNHLEGARE 1031
DB 1005 SNHMSREE 1014

RESULT 7
US-07-743-357-9
; Sequence 9, Application US/07743357
; Patent No. 5858646
; GENERAL INFORMATION:
; APPLICANT: Kang, Yong C.
; TITLE OF INVENTION: Polypeptide having immunological
; TITLE OF INVENTION: activity for use as diagnostic reagent and/or vaccine
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KIRBY EADES GALE BAKER
; STREET: Box 3432, Station D
; CITY: Ottawa
; STATE: Ontario
; COUNTRY: Canada
; ZIP: K1M 1H8
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07743,357
; FILING DATE: 21-AUG-1991
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/CA90/00062
; FILING DATE: 23-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Gale, Edwin J.
; REGISTRATION NUMBER: 28,584
; REFERENCE/DOCKET NUMBER: 30924-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 237-6900
; TELEFAX: (613) 237-0045
```

```
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1003 amino acids
; TYPE: amino acid
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Human immunodeficiency virus type 1
; STRAIN: MAL
; US-07-743-357-9
```

```
Query Match 57.9%; Score 3155; DB 2; Length 1003;
Best Local Similarity 57.3%; Pred. No. 5.2e-261;
Matches 576; Conservative 177; Mismatches 236; Indels 16; Gaps 5;

QY 32 PGSSSGSTGEIYAAREKTERAERTIQGSDRGLTAPRAGGDTIQAGTNRGLAA---PQF 88
DB 9 PQGKAREFPSEOTRANSPTSRELR--VMGDKTLE-----TGAERQGVISFSPQI 58
QY 89 SLWKRPVVTAYIEGQPVVELLDTGADDSIVAGIELGNYSPTKIVGGIGGFINTKEVKNVE 148
DB 59 TLMQRPVVTYVRVGQQLKEALLDTGADDTVLEEINLPKWKPMIGIGGFYKVRQDQIL 118
QY 149 IEVLNKKVRATIMTGDPTINIFGRNLTALGMSLNLPAKVEPIKIMLPGKDGPKLRQW 208
DB 119 IEICGKKAICTILVGTPTVNIIGRNMLTQIGCTLNFPIETVPVKLPGMDGPRVKQW 178
QY 209 PLTKIEKALKEICERKEGQLEAPPTNPYPTFAIKKKDKNKKWMLIDPRELNKVT 268
DB 179 PLTEEEKIKALTEICKOMEKEGKILKIGPENPYPTVFAIKKKDKSTKWKLVNPRELNKRT 238
QY 269 QDTEIOLGIPHPAGLAKRRITVLDVGDAYSILHEDPRVPTAETLPSVNNAPGKRY 328
DB 239 QDFEWQLGIPHPAGLAKRRKSVTVLDVGDAYSFVLDDEFKRYTATFIPISINNETPGIRY 298
QY 329 IYKVLPGQWKGSPAIFQHTMRQVLEPFRKANKDVIITQYMDLIIASDRTDLEHDRVVLQ 388
DB 299 QYNVLPQWKGSPAIFQSSMTKILEPFTKNPEIVYQYMDLIVGSDLEIGQHRKIEE 358
QY 389 LKELLNGLGFSPDEKFDQDPYHWMGYELWPTKWLQIKLPQKEIWTVNDIQKLVGL 448
DB 359 LREHLKLGWFTTDDKKHQKEPPFLWGMGYELHDPKWTVPQIQLPKDESMTVNDIQKLVGL 418
QY 449 NMAAQLYPGIKTKHLCLIRGKMTLTEEQVWTELAELAEENRIILSQEQEGHYQEEKE 508
DB 419 NMAQIYPGIKVQLCKLKGAKALTDIYPLTAEAELEAENREILKEPVHGVYIDPSKD 478
QY 509 LEATVQDQENQWYTKIHQEE-KILKVGKYAKVKNPHTNGIRLLAQVQVQKIGKEALVWG 567
DB 479 LIAEIQKQGGQWYQIYQEQYKLNKTGKVARIKSAHTNDVKQLTEAVQIAQESIVWG 538
QY 568 RPKFHLPVREIWEQWMDNYQVWTWIPDWDVSTPPLVRLAFNLVGDPIPGATFTYD 627
DB 539 KTPKRLPQKETWEAWTEYQWATWPEWFEVNTPLVKLWYQLETEPTEIYGAETFYVDG 598
QY 628 SCNRQSEKAGVYTDGRGDKVKKLPQTTNQAAELAEAFAMALTDSPGKVINIIVDSQVVMG 687
DB 599 AANRETKKAGVYTDGRGDKVKKLPQTTNQAAELAEAFAMALTDSPGKVINIIVDSQVVMG 658
QY 688 ISASQPTSESKIVNOIIEEMIKKEIYVAVWPAHKGIGGNQEVHDLVSGQIRQVLFLEK 747
DB 659 LIAQAPDKSESEIVNQLIEQLQKDKVYLSWPAHKGIGGNQEVHDLVSGQIRQVLFLEK 718
QY 748 IEPAQEEHEKYHSNVELSHKFGIPNLVARQIVNSCAQCKQGEATHGQVNAELGTWQMD 807
DB 719 IDKAQEEHEKYHSNWRAMASDENLPPIVAKEIVASCDKQCKGAEAHGQVDCSPGWLQD 778
QY 808 CTHLECKIIVAVHIVASGFEAEVIPSQRQTAELFLKLKASRPITHLHTDNCANTSO 867
DB 779 CTHLECKIIVAVHIVASGFEAEVIPSQRQTAELFLKLKASRPITHLHTDNCANTSO 838
```



```

; ATTORNEY/AGENT INFORMATION:
; NAME: Gale, Edwin J.
; REGISTRATION NUMBER: 28,584
; REFERENCE/DOCKET NUMBER: 30924-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 237-6900
; TELEFAX: (613) 237-0045
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1004 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Human immunodeficiency virus type 1
; STRAIN: SF2
; US-07-743-357-7

Query Match          57.2%; Score 3117; DB 2: Length 1004;
Best Local Similarity 56.8%; Pred. No. 6.3e-258;
Matches 567; Conservative 179; Mismatches 241; Indels 12; Gaps 6;

QY 38 GSTGEIYAAREKTERAETIQ--GSDRLGTAPRAGGDTIQGATNRGLAAPQFSLWKRP 94
DB 11 GKAREFSSEQTRANSPTRELOVWGGNNLSL--EAGADR-QCTVSNF--PQITLWQRP 65
QY 95 VVTAYTEGQPVVLLDTGADDSIVAGIELGNYSKPKIVGGIGGIFINTKKNVEIEVLNK 154
DB 66 LVTRIGGQKLEALLDTGADDTVLEENMLPGKWKRMIGGIGFIVKRYQDQIPVETCGH 125
QY 155 KVRATIMTGTPTNIFGRNLTALGMSLNPVAKVEPIKIMLPGKDGPKLRQMPLTKEK 214
DB 126 KATGVLVGTPTNIGRNLLTQIGCTLPPIETVPVKLPKGMGDPKVKOMPLTEK 185
QY 215 IEALKEICEKMEKEGOLLEAPPTNPNYTPPTFAIKKDKNKRMLIDFRELNKVTDQFTEI 274
DB 186 IKALVEICTEMEKEGKISKIGPENPTYVFAIKKDKSTKRWKLVDRELNKRKTQDFWEV 245
QY 275 QLGIPHPAGLAKRRITVLDVGDAYSIPLHEDFRPYTAFTLSVNNAPGKGIYKVLPL 334
DB 246 QLGIPHPAGLKKKSVTLVDVGDAYSFVPLDKDFRYTAFTIPSINNETPGIRYQYNLPL 305
QY 335 QGKKGSPAIPOHTMROVLEPRKANKDVIIIOYMDLILASDRDTLDEHDRVVLQKLELN 394
DB 306 QGKKGSPAIPOSSMTKILEPFRKQNDIVIIYQTMDDLYGSDLEIGQHRTKIEELRQHL 365
QY 395 GLGFSTPDEKFPQPPYHMGYELWPTKWKQLQPLQKBEIWTVDIQKLVGLVNWAAQL 454
DB 366 RWGFTTPDKKHQKPEPPFLMWGYELHPDKWTVPIMLPKEDSVTVNDIQKLVGLNWSQI 425
QY 455 YPGIKYKHLRLTRGKMTTEEVQWTELAELEENRIILSQBQEGHYOYEKELEATVQ 514
DB 426 YAGIKYKQLKLLRGTKALTEVPLTEAELEAENREILKEPVHEVYVDPSSKDLVAEIQ 485
QY 515 KQENQWTKYHOEE-KILKVGKYAKVKNTHNGIRLLAQVYOKICKKEALVIVGRPKPH 573
DB 486 KQCGQQTVOIYEPFNKLTGKYGARMGAHTNDVKOLTEAVQKVSTESIVINGKIPKPK 545
QY 574 LPVERIEWQWNYQVWTIPDQWDFVSTPPLVRLAFNLVGDPIPGAETFTYDGCNROS 633
DB 546 LPIQKETWEAWWYQATWIPWEFVNTPLPKLWYQLEKEPIVGAETFYVDGAANRET 605
QY 634 KEGKAGYVTRDGKDKVKLEQTTNQQAELAEAFAMALTDGPKVNIIVDSQYVNGISASQP 693
DB 606 KLGKAGYVTRDGKQVSVSADTTNQTKELQAIHALQDSGLEVNIIVTDSQYALGIQAQP 665
QY 694 TSESKIVNQIIEEMIKKAIYVAVVPAHKGGIGGNOVDHLVSQGRVOLFLEKIEPAQE 753
DB 666 DKSESELSQIIBQLIKKVKYLANVPAHKGGIGGNOVDKLVASAGIRKVLFLNGIDKAE 725
```

```

RESULT 10
US-07-743-357-1
; Sequence 1, Application US/07743357
; Patent No. 5858646
; GENERAL INFORMATION:
; APPLICANT: Kang, Yong C.
; TITLE OF INVENTION: Polypeptide having immunological
; TITLE OF INVENTION: activity for use as diagnostic reagent and/or vaccine
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KIRBY EADES GALE BAKER
; STREET: Box 3432, Station D
; CITY: Ottawa
; STATE: Ontario
; COUNTRY: Canada
; ZIP: K1M 1H8
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/743,357
; FILING DATE: 21-AUG-1991
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/CA90/00062
; FILING DATE: 23-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Gale, Edwin J.
; REGISTRATION NUMBER: 28,584
; REFERENCE/DOCKET NUMBER: 30924-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 237-6900
; TELEFAX: (613) 237-0045
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1005 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Human immunodeficiency virus type 1
; STRAIN: HXB2
; US-07-743-357-1

Query Match          57.2%; Score 3117; DB 2: Length 1005;
Best Local Similarity 56.8%; Pred. No. 9.4e-258;
Matches 564; Conservative 178; Mismatches 248; Indels 6; Gaps 3;
```

QY	38	GSTGEIYAAAREKTERAERETIQGSDRGLTAPRAGGDTIQGATNRGLAAPQFSLWKRPVVT	97
Db	12	GRAREESSQTRANSPTRELQWGDNNSPSEAGADRGTVSFNF--FOVTLWQRLPVT	69
QY	98	AYIEGQPVVLLDGTGADDSIVAGIELGNNSPKIVGGIGGFINTKEYKNVEIEVLNKKYR	157
Db	70	IKIGQLKEALDGTADDTVLEEMSLPGRWKPKMIGGIGGFVKVQYDQILIEICGHKAI	129
QY	158	ATIMTGDPTINIFGRNLTALGMSNLNLPVAKVEPIKIMLPGKDGPKLQWPLTRKIEA	217
Db	130	GTVLVGPPTVNIIGRNLLTQIGCTLNFPSPIETVPVCLKPGMDGPKVKQWPLTEKIRA	189
QY	218	LKEICEKMEKEGLEEAPPNTNPTFAIKKDKKNKRWMLDFRELNNKVTQDFTRIQLG	277
Db	190	LVEICTEMEKEGKISKIGPENYPNPVFAIKKDKSTKRWKLVDFRELNKRTQDFEVLQ	249
QY	278	IPHPAGLAKKRITVLDVGDAFYSIPLHEDFRPYTAFTLPSVNNAPGKRYIKYVLPOGW	337
Db	250	IPHPAGLKKKSVTVLDVGDAFYSVPLDEDFRKYTAFTTIPSINNETPGIRYQYNVLPQGW	309
QY	338	KGSPAIFQHTMRQVLEPPFRKANKDVIIIOYMDLIIASDRTDLEHDRVVLQKELLNGLG	397
Db	310	KGSPAIFQSSMTKILEPPFRKQPDVIVIOYMDLVGSDLEIGQHRTKIEELRQHLRWG	369
QY	398	FSTPDEKFOKPPYHWMGYELWPTKWKLOKIQLPQKEIWTVNDIQKLVGLNWAQOLYPG	457
Db	370	LTPDKKHQKPPFLWMGYELHPDKWTQVPIVLPKDSWTVNDIQKLVGLNWAQOLYPG	429
QY	458	IKTKHLCLRLIRKMTLITEVQVTELAELAELENRRILSQBOGHHYQOEKELEAVVKDQ	517
Db	430	IKVRQLCKLLRTKALTEVIPTEAELELAENREILKPEVHGVIYDPSKDLIAEQKQ	489
QY	518	ENQWYKIHQEE-KILKVGKAKVKNTHNGIRLLAQQVVKIGKEALVINGRIPFHLPV	576
Db	490	QGWYQIYQEFKMLTKGYARMGAHTNDVKQTEAVQKLTTSIVINGKTPFKLPI	549
QY	577	EREINQWMDNVTQWTVTPDMDFVSTPLVRLAFNLGDPPIGAETFTYDSCNRSKRG	636
Db	550	OKETWETWTEYQATWPEWEFVNTPLVLKWLWYQLEKEPIVGAETFFYVDGAANRETKLG	609
QY	637	KAGYVTDGKDKVKKLEQTNOQAELEAFAMALTDGSGPKVNIIVDSQYVMGISAQPTES	696
Db	610	KAGYVTRNGRQKVVTLTDTTNOKTELOAIYLALQDSGLEVNIVTDSQYALGIIQAQPDG	669
QY	697	ESKIVNOIIEEMIKKEATIVAVPAHKGIGGNQEVDLVSGIROVLPLEKTEPAQEEHE	756
Db	670	ESELVNOIIEQILKEKYLAWPAHKGIGGNEQDKLVSGIRKVLFDGIDKQADEHE	729
QY	757	KYHSNVKELSHKFGIPNLVARQVNSCAQCOQKGEAIIHQVNAELGTWQMDCTHLEGKII	816
Db	730	KYHSNWRAMASDFNLPPVVAKEIVASCDKQLKGEAMHGOVDCSPGIMQLDCTHLEGKI	789
QY	817	IVAVHVASGFTEAEVIPAQESGQTALFLKLKASRWPITHLTDNGANFTSQBVKVAWVI	876
Db	790	LVAHVASGYIEAEVIPAETQETAYFLKLKLAGRPVKTITHDNGSNFTGATVRAACWMA	849
QY	877	GIEQSGFVPYNPQSGVVEAMHHLKNQISIREQANTTETIVLMAIHCNFKRGGIGD	936
Db	850	GIKQEGIPYNPQSGVVESNMKELKKIIGQVRODAEHLKTAVQMAVFHNFKRGGIGG	909
QY	937	MTPSERLINMITTEIQFLQAKNSKLDKDFRVYFREGRDQLWKGPGLWKGEQAVLVKV	996
Db	910	YSAGERIVDIIATDITQKLOKQITKIQNFRVYVYRNSRLWKGPAKLWKGEQAVVIQD	969
QY	997	CTDIKILPRKAKIIRDYGGREMDSGSHLEGARED 1032	
Db	970	NSDIKVPRRKAKIIRDYGKQM---AGDDCVASRQD 1002	
RESULT 11			
US-07-743-357-3			
; Sequence 3, Application US/07743357			

;	Patent No. 5858646		
;	GENERAL INFORMATION:		
;	APPLICANT: Kang, Yong C.		
;	TITLE OF INVENTION: Polypeptide having immunological		
;	activity for use as diagnostic reagent and/or vaccine		
;	NUMBER OF SEQUENCES: 22		
;	CORRESPONDENCE ADDRESS:		
;	ADDRESSEE: KIRBY EADES GALE BAKER		
;	STREET: Box 3432, Station D		
;	CITY: Ottawa		
;	STATE: Ontario		
;	COUNTRY: Canada		
;	ZIP: K1M 1H8		
;	COMPUTER READABLE FORM:		
;	COMPUTER: IBM PC compatible		
;	OPERATING SYSTEM: PC-DOS/MS-DOS		
;	SOFTWARE: PatentIn Release #1.0, Version #1.30		
;	CURRENT APPLICATION DATA:		
;	APPLICATION NUMBER: US/07/743,357		
;	FILING DATE: 21-AUG-1991		
;	CLASSIFICATION: 424		
;	PRIOR APPLICATION DATA:		
;	APPLICATION NUMBER: PCT/CA90/00062		
;	FILING DATE: 23-FEB-1990		
;	ATTORNEY/AGENT INFORMATION:		
;	NAME: Gale, Edwin J.		
;	REGISTRATION NUMBER: 28,584		
;	REFERENCE/DOCKET NUMBER: 30924-2		
;	TELECOMMUNICATION INFORMATION:		
;	TELEPHONE: (613) 237-6900		
;	TELEFAX: (613) 237-0045		
;	INFORMATION FOR SEQ ID NO: 3:		
;	SEQUENCE CHARACTERISTICS:		
;	LENGTH: 1016 amino acids		
;	TYPE: amino acid		
;	STRANDEDNESS: not relevant		
;	TOPOLOGY: not relevant		
;	MOLECULE TYPE: protein		
;	HYPOTHETICAL: NO		
;	FRAGMENT TYPE: Internal		
;	ORIGINAL SOURCE:		
;	ORGANISM: Human immunodeficiency virus type 1		
;	STRAIN: BH5		
;	US-07-743-357-3		
;	Query Match	57.2%;	Score 3116; DB 2; Length 1016;
;	Best Local Similarity	56.3%;	Pred. No. 1.2e-257;
;	Matches	564;	Conservative 182; Mismatches 249; Indels 6; Gaps 3;
QY	33	SGSSSGSTGEIYAAAREKTERAERETIQGSDRGLTAPRAGGDTIQGATNRGLAAPQFSLWK	92
Db	18	SQTRANSPTISSEQTRANSPTRELQWGDNNSPSEAGADRGTVSFNF--FOVTLWQ	75
QY	93	RPVVTAYIEGQPVVLLDGTGADDSIVAGIELGNNSPKIVGGIGGFINTKEYKNVEIEVL	152
Db	76	RPLVTRIKIGQLKEALDGTADDTVLEEMSLPGRWKPKMIGGIGGFVKVQYDQILIEIC	135
QY	153	NKKVRATWTGPTPINIFGRNLTALGMSNLNLPVAKVEPIKIMLPGKDGPKLQWPLTK	212
Db	136	GHKALGTVLVGPPTVNIIGRNLLTQIGCTLNFPSPIETVPVCLKPGMDGPKVKQWPLTE	195
QY	213	EKIEALKEICEKMEKEGLEEAPPNTNPTFAIKKDKKNKRWMLDFRELNNKVTQDFT	272
Db	196	EKIKALVEICTEMEKEGKISKIGPENYPNPVFAIKKDKSTKRWKLVDFRELNRRTQDFW	255
QY	273	ETQLGIPHPAGLAKKRITVLDVGDAFYSIPLHEDFRPYTAFTLPSVNNAPGKRYIKYV	332
Db	256	EVLQGIPHPAGLKKKSVTVLDVGDAFYSVPLDEDFRKYTAFTTIPSINNETPGSYQYNV	315
QY	333	LPQWKGSPAIFQHTMRQVLEPPFRKANKDVIIIOYMDLIIASDRTDLEHDRVVLQKEL	392
Db	316	LPQWKGSPAIFQSSMTKILEPPFRKQPDVIVIOYMDLVGSDLEIGQHRTKIEELRQH	375

APPLICANT: CHANG, Nancy T.
APPLICANT: GALLO, Robert C.
APPLICANT: WONG-STAL, Flossie
TITLE OF INVENTION: CLONING AND EXPRESSION OF HTLV-III DNA
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10154-0053

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/463.210

FILING DATE: 05-JUN-1995

CLASSIFICATION: 436

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 06/693,866

FILING DATE: 23-JAN-1985

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 06/659,339

FILING DATE: 10-OCT-1984

ATTORNEY/AGENT INFORMATION:

NAME: Serunian, Leslie A.

REGISTRATION NUMBER: 35,353

REFERENCE/DOCKET NUMBER: 2026-4193US2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800

TELEFAX: (212) 751-6849

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 1015 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORGANISM: HTLV-III

FEATURE:

NAME/KEY: Protein

LOCATION: 1..1015

OTHER INFORMATION: /note= "pol protein of HTLV-III"

US-08-463-210-9

Query Match

Best Local Similarity 57.1%; Score 3112; DB 3; Length 1015;

Matches 563; Conservative 183; Mismatches 249; Indels 6; Gaps 3;

33 SGSSGSGTGEIYAAREKTERAERTIQGSDRGLTAPRAGGDTIQGATNRGLAAPQFSLWK 92

18 SEQTRANSPITISSEQTRANSPTRRELQVGRDNNSPSEAGADRGQTVSFNF--PQITLWQ 75

93 RPVTAYTEGQVPEVLLDTGADDSIVAGIELGNYSPIKIVGIGGIFINTKEYKNVEIEVL 152

76 RPLVTIKIGIKKEALLTGADDTVLEMSLPGWRKPKMIGGIGGIFKVRQYDQILIEIC 135

153 NKKVRATIMTGDPTINIFGRNLTALGMSLNPVAKVPEPIKMLKPGKDGPKLRQWPLTK 212

136 GHKAGTVLVGPTVPVNIIGRNLLTOIGCTLNFPISPIETVPVKLAPGMDGPKVKQWPLTE 195

213 EKIEALKICEKMEKEGEEAPPNTPTFAIKKKDKNRMLIDFRELNKVTDQFT 272

196 EKIKALVEICTEKEGKISKIGPENPYTPVFAIKKKDKNRMLIDFRELNKVTDQFW 255

273 EIQIGIPHAGLAKRRITVLVDGDAYESIPLHEDFRPYTAFTLPSVNNAPGKRYIKV 332

256 EVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIPSINNETPGIRYQYNV 315

333 LPQGWKGSIPAIFQHTMROVLPERKANKDVIIIOYMDLIIASDRDLEHDRVVLQKL 392

336 LPQGWKGSIPAIFQSSWTKILEPFFKQNDIVIIQYMDLIIYVGSDEILGQHRTKIEELQ 375

393 LNLGLFSTPDEKQFQPPYHMMGYELWPTKWKLOKIQLPQKEIWTWNIDIKLVGLNMAA 452

376 LLRWGLTTPDKKKHQKPEPPFLWNGYELHLPDKWTVPQVPLPEKDSWTVNDIQLVGLNMA 435

453 QLYPGIKTKHLCLRLIRGKMTLFEVOWTELAELAEENRIILISQEQEGHYHYYEELAT 512

436 QIYPGKVRQCKLLRGTKALTEVPLTEAELEAENREILKEPVGHVYDPSDLIAE 495

513 VOKDOENQWYKIHQEE-KILKVGKYAKVKNTHNGIRLLAOWGVKIGKEALVIGRIPK 571

496 IQKQGGQWTVIYQEPFKNLKTGYARMGAHTNDVKQLTAVOKITTESIVINWKTPK 555

572 PHLVERIEWQWMDNYQWTVIPDWFVSTPPLVRLAFNLVGDPIPGAETFTDSCNR 631

556 FKLPIQKETWETWTEYQATWPEWEFVNTPLVKLWQLEKEPIVGAETFTVDGAANR 615

632 QSKCKAGYVTRDGRKDKVKKLEQTTNQQAELAFAMALTDSPKVNIIIVDSQYVNGIS 691

616 ETKLGAGYVTKNGKQVYVPLTNTNQKTEQAIIYALALODSGLEVINIVTDSOYALGITA 675

692 QPTESEKIVNOIIEEMIKKEAIIYVAVWPAHKGIGGNOEVDHLVSGIGROVLFLKIEPA 751

676 QPKSESELVNOIIEOLIKKEKYLAWPAHKGIGGNEQVQKLVASGIRKILFLDGDKA 735

752 QEEHEKYSNVKELSHKFGIPNLVARQIVNSCAQCOQKEATHGOVNAELGTWQMDCTHL 811

736 QDEHEKYSNMRAMASDENLPPVAKETIVASCDCQLKGEAMHGQVDCSPGIWQDCTHL 795

812 ECKIIIVAVHVASGIEAEVIPQESGRQTALELLKLSRWPTIHLHTDNGANFTSOEYKM 871

796 ECKVILVAVHVASGIEAEVIPAETGOETAYELLKLAGWPVKTIHTDNGSNFTSATVKA 855

872 VAWWIGIEQSGVYPNPOSGVVEAMNHLKNOISIRIEQANTITETIVLMAHCHNFKRR 931

856 ACWAGIKQEGFIPYNPOSGVVEAMNHLKNOISIRIEQANTITETIVLMAHCHNFKRR 915

932 GGIGMDTPSERLINMITTEQETQFLOAKNSKLUKDFRVYFREGROQLWKGPGELLWGEA 991

916 GGIGYSAGERIVDIATDIQTKELQKQITKIQNFVRYVYRDNPLWKGPAKLLWGEA 975

992 VLVKVTGDIKIIPRKAIIIRDYGGQEMDSSHLEGARE 1032

976 VVIQNDSDIKVYVPRRKAIIIRDYGKQ---AGDDCVASQD 1013

RESULT 14

US-09-124-900-3

Sequence 3, Application US/09124900

Patent No. 6268484

GENERAL INFORMATION:

APPLICANT: KATINGER, Hermann

APPLICANT: BUCHACHER, Andrea

APPLICANT: ERNST, Wolfgang

APPLICANT: BALLAUN, Claudia

APPLICANT: PURTSCHER, Martin

APPLICANT: TRKOLA, Alexandra

APPLICANT: PREDL, Renate

APPLICANT: SCHMATZ, Christine

APPLICANT: KLINA, Annelies

APPLICANT: STEINDL, Franz

APPLICANT: MUSTER, Thomas

TITLE OF INVENTION: HIV-Vaccines

FILE REFERENCE: 1939-112p

CURRENT APPLICATION NUMBER: US/09/124,900

CURRENT FILING DATE: 1998-07-30

PRIOR APPLICATION NUMBER: PCT/EP95/01481

PRIOR FILING DATE: 1995-04-19

NUMBER OF SEQ ID NOS: 11

; SOFTWARE: Patentin version 3.0		
; SEQ ID NO 3		
; LENGTH: 1015		
; TYPE: PR		
; ORGANISM: Human immunodeficiency virus type 1		
US-09-124-900-3		
Query Match		57.1%; Score 3112; DB 3; Length 1015;
Best Local Similarity		56.2%; Pred. No. 2.6e-257;
Matches	563; Conservative	183; Mismatches 249; Indels 6; Gaps 3;
Qy	33	SGSSGSGTGEIYAAREKTERAERETIQGSDRGLTAPRAGDGTIQGATNRGLAAPQFSLWK 92
Db	18	SEQTRANSPITSSQTRANSPTRELQVWGRDNNSPEAGADRGQTVSENF--PQITLWQ 75
Qy	93	RPVVTAYIEGQPVVELDGTGADDSIVAGIELGNNSPKIVGGTGGFTNTEYKNVEIEVL 152
Db	76	RPLVTIKIGGQLEKALDGTGADDTVLEEMSLPGRWKPKMIGGIGGFTKVRQYQILIEIC 135
Qy	153	NKVRATIMTGDPTINIFGRNLTALGMSLNPVAKVEPIKMLKPGKDGPKLRQWPLTK 212
Db	136	GKAIQVILVGPTEVNIIGRNLLTQIGCTLNFPISPIETVPVKLPGMDGPKVQWPLTE 195
Qy	213	EKIEALKEICEKMEKEQLEAEAPTNPYNTPTFAIKKKDKNKRMLIDFRELNKVTDFT 272
Db	196	EKIKALVEICTEMEKEGKISKIGPENPYNPVFAIKKKDKSTKWRKLVDFRELNKRTQDFW 255
Qy	273	EIQIGIPHAGLAKKRITVLDVGDAYSIFLPHEDFRPYTAFTLPSVNNAEPEKRYIKV 332
Db	256	EVQIGIPHAGLAKKRITVLDVGDAYSIFLPHEDFRPYTAFTLPSVNNAEPEKRYIKV 315
Qy	333	LPGWKGSPALFOSSMTKILEPEKKQNPDIIVIQYMDLVGSDLEIGQHRTKIEELRQH 375
Db	316	LPGWKGSPALFOSSMTKILEPEKKQNPDIIVIQYMDLVGSDLEIGQHRTKIEELRQH 375
Qy	393	LNLGFTSPDEKQKOPPYHWMGYELMPTKWKLOKQOLPOKEITWVNDIQKLVNLWAA 452
Db	376	LLRWGLTPDKKHQKPEPFLMWMGYELMPTKWKLOKQOLPOKEITWVNDIQKLVNLWAA 435
Qy	453	QLYPGIKTKHLRLIRKMTLTEVOWTELAELAENRILISQEQEGHYQEKELEAT 512
Db	436	QIYPGIVKQRLKRLTKALTEVPLTEBAELEAENRILKPEVHGVIYDYSKDLIAE 495
Qy	513	VQDOENQWYTKIHQEE-KILYGVKAKVKNTHNGIRLLAQVQVQKIGKALYIWRGPK 571
Db	496	IQXGOGOWTYQIYQEPFKNLTKGYARMRGAHTNDVKQLTEAVQKTTESIVIGWTKPK 555
Qy	572	FHLPVERIEQWQNDWYQVTPDWDVSTPLVLRLAFNLVGDPIPGAETFTYDGSNR 631
Db	556	FKLPQIKETWETWTEYQATWIPWEPFVNTPLVLWYQLEKEPIVGAETFTYVDGAANR 615
Qy	632	QSKEGKAGYVTRGKDKVKLEQTTNQAELAFAFAMALTDGSPKVIIVDSQVVMGISAS 691
Db	616	ETKLGRAGYVTKGRQVWPLTNTWTKTELQAIYALQDSGLEVNIVTDSQVAGLIQA 675
Qy	692	OPTSESEKYNQIIEEMIKKEATYVAVWPAHKGIGGQNEVDHLVSGIQVLFLEKIEPA 751
Db	676	QPKSESELVNIIEQLIKKEKVLAWVPAHKGIGGQNEVDHLVSGIQVLFLEKIEPA 735
Qy	752	QEBEYKHSNVELSHKFGIPNLVARIQVNSCAQCOQKGEAIIHGQVNAELGTWQMDCTHL 811
Db	736	QEBEYKHSNWRAMASDFNPPVYAKEIVASCDQLKGEAMHGQVDCSPGIMQDLCTHL 795
Qy	812	EGRIIVAVHVASGTEAEVPOESGROATFLKLKASRMPITHLTDNGANTTSQEVKM 871
Db	796	EGKIVLVAVHVASGYEAEVPAETQETAYFLKLKAGRPVTKITHDNGSNFTSATVKA 855
Qy	872	VANWIGIEQSGFYVPYQSQGVYVAMNHLKNOISRIEQAQNTIETIYVMAIHCMPKRR 931
Db	856	ACWAGIKQEBFGIPYQSQGVYVAMNHLKNOISRIEQAQNTIETIYVMAIHCMPKRR 915
Qy	932	GGTGMTPSRLINMTTEQIEQIFLOAKNSKLDFRYVYFREGRDQLWKGCELLWKGEA 991

Db	916	GGIGYSGASRIVDIIATDIOTKRELQKQITKIQNFVYVYRDSRNPLWKGPAKLLWKGEA 975
Qy	992	VLVKGTDIKIIPRRAKAIIRDYGGQEMDMSGHLEGARED 1032
Db	976	VVIQDNSDIKVPVRRKAIIRDYQKQ---AGDDCVASROD 1013
RESULT 15		
US-07-743-357-2		
; Sequence 2, Application US/07743357		
; Patent No. 5858646		
; GENERAL INFORMATION:		
; APPLICANT: Kang, Yong C.		
; TITLE OF INVENTION: Polypeptide having immunological		
; TITLE OF INVENTION: activity for use as diagnostic reagent and/or vaccine		
; NUMBER OF SEQUENCES: 22		
; CORRESPONDENCE ADDRESS:		
; ADDRESSEE: KIRBY EADES GALE BAKER		
; STREET: Box 3432, Station D		
; CITY: Ottawa		
; STATE: Ontario		
; COUNTRY: Canada		
; ZIP: KIM 1H8		
; COMPUTER READABLE FORM:		
; MEDIUM TYPE: Floppy disk		
; COMPUTER: IBM PC compatible		
; OPERATING SYSTEM: PC-DOS/MS-DOS		
; SOFTWARE: Patentin Release #1.0, Version #1.30		
; CURRENT APPLICATION DATA:		
; APPLICATION NUMBER: US/07/743,357		
; FILING DATE: 21-AUG-1991		
; CLASSIFICATION: 424		
; PRIOR APPLICATION DATA:		
; APPLICATION NUMBER: PCT/CA90/00062		
; FILING DATE: 23-FEB-1990		
; ATTORNEY/AGENT INFORMATION:		
; NAME: Gale, Edwin J.		
; REGISTRATION NUMBER: 28,584		
; REFERENCE/DOCKET NUMBER: 30924-2		
; TELECOMMUNICATION INFORMATION:		
; TELEPHONE: (613) 237-6900		
; TELEFAX: (613) 237-0045		
; INFORMATION FOR SEQ ID NO: 2:		
; SEQUENCE CHARACTERISTICS:		
; LENGTH: 1016 amino acids		
; TYPE: amino acid		
; STRANDEDNESS: not relevant		
; TOPOLOGY: not relevant		
; MOLECULE TYPE: protein		
; HYPOTHETICAL: NO		
; FRAGMENT TYPE: internal		
; ORIGINAL SOURCE:		
; ORGANISM: Human immunodeficiency virus type 1		
; STRAIN: BH102		
US-07-743-357-2		
Query Match		57.1%; Score 3112; DB 2; Length 1016;
Best Local Similarity		56.2%; Pred. No. 2.6e-257;
Matches	563; Conservative	183; Mismatches 249; Indels 6; Gaps 3;
Qy	33	SGSSGSGTGEIYAAREKTERAERETIQGSDRGLTAPRAGDGTIQGATNRGLAAPQFSLWK 92
Db	18	SEQTRANSPITSSQTRANSPTRELQVWGRDNNSPEAGADRGQTVSENF--PQITLWQ 75
Qy	93	RPVVTAYIEGQPVVELDGTGADDSIVAGIELGNNSPKIVGGTGGFTNTEYKNVEIEVL 152
Db	76	RPLVTIKIGGQLEKALDGTGADDTVLEEMSLPGRWKPKMIGGIGGFTKVRQYQILIEIC 135
Qy	153	NKVRATIMTGDPTINIFGRNLTALGMSLNPVAKVEPIKMLKPGKDGPKLRQWPLTK 212
Db	136	GKAIQVILVGPTEVNIIGRNLLTQIGCTLNFPISPIETVPVKLPGMDGPKVQWPLTE 195
Qy	213	EKIEALKEICEKMEKEQLEAEAPTNPYNTPTFAIKKKDKNKRMLIDFRELNKVTDFT 272


```
Db 196 ERIKALVEICTEMEKEGKISKIPENYNTPTFAIKKDDSTKRWKLVDPFRELNKRQTDFW 255
QY 273 EIQLGIPHPAGLAKRRITVLDVGDAYSIPLHEDFRPYTAFTLPSVNNAEFGKRYIKV 332
Db 256 EVQLGIPHPAGLKKKSVTLVDGDAYFSVPLDEDFRKYTAFTIPSINNETPGIRYQNV 315
QY 333 LPQGWKGSIPAIFQHTMRQVLEPFRKANKDVIIIOYMDDLIASDRTDLEHDRVVLQKL 392
Db 316 LPQGWKGSIPAIFQSSMTKILEPFRKQNPDIYQYMDLVGSDLEIGQHRTKIELRQH 375
QY 393 LNLGFSPPDEFKQDPYPHMGYELWPTKWKQLQIPLQKQIWTVDIQLVGLVNWAA 452
Db 376 LLRWGLTTPDKKHQKEPPLWNGYELHPDKWTVQVPLPEKDSWTVNDIQLVGLNWA 435
QY 453 QLYPGIKYKHLRLRGKMTLFEVQWTELAELAEENRIILSOQEGHYOEEKELEAT 512
Db 436 QIYPGIKYKQLCKLLRGKALTEVPLTEEALELAENREILKEPVHGVYDPSKDLIAE 495
QY 513 VOKQOENQWYKIHQEE-KILKVGKYAKVKNTHNGIRLLAQVQKIGKEALVINGRIPK 571
Db 496 IQKQOGQWYIYQEPKLNKTGKYARMGAHTNDVKLTEAVQKITTESIVINGKTPK 555
QY 572 PHLPVEREIQWQWNYQVTVIPDWDVFSVTPPLVRLAFNLVGDPIGAETFTDGSNCR 631
Db 556 FKLPIQKETWETWYQATWIPWEFVNTPLVKLWYQLEKEPIVGAETFTYVDAANR 615
QY 632 OSKEGKAGVYDGRGKVKLQTTNQOAELEAFAMALTDGPKYNIIVDSQYVNGISAS 691
Db 616 ETKLGKGVYTNKGKQKVVPLNTNTNQKTELOAIYALQDSGLEVNIIVDSYALGIQA 675
QY 692 QPTSESKIVNQIIEEMIKKEAIYVAVYPAHKGIGGNOEVHLSQIGROVLFLEKIPPA 751
Db 676 QPDKSESELVNIQIQLKKEKYLAWYPAHKGIGGNEQVQKLVGAGIKRILFLDGIDKA 735
QY 752 QEEHEKYSNVELSHKFGIPNLVARQIVNSCAQOQKGEATHGQVNAELGTWQMDCTHL 811
Db 736 ODEHEKYSNWRAMASDENLPPVAKEIVASCDKQKLGAEAMHGQVDCSPGIWQDCTHL 795
QY 812 EKKIIVAVHVASGFIIEAIVPOESSGROTALFLKLASRPWITHLTONGANFTSQEYKM 871
Db 796 EKGKIVAVHVASGFIIEAIVPAETQOETAYELLKLAGRPVKTILHTDNGSNFTSATVKA 855
QY 872 VAWTIGIEQSGFYNPYNSQGVVEAMNHLKNOISRIREQANTIIETIVLMAIHCNFKRR 931
Db 856 ACWAGIQRQEGIPYNSQGVVESNMKELKKIIGVRDQAEHLATATQMAVFIHFKRK 915
QY 932 GGIGDMTYSERLINMITTEQETQFLOAKNSKLKDPFRVYFREGRODLWKGPGELLWKGE 991
Db 916 GGIGYSAGERIVDIATDIQTKELQKQITKIQNFRVYRDSRNPRLWKGPAKLLWKGE 975
QY 992 VLVKVGTDIKIIPRKAIIIRDYGGQEMDSGSHLEGARE 1032
Db 976 VVIQNSDLKVYPRKAKIIRDYGKQ---AGDDCVASRQD 1013
```

RESULT 16

```
US-07-743-357-5
; Sequence 5, Application US/07743357
; Patent No. 5858646
; GENERAL INFORMATION:
; APPLICANT: Kang, Yong C.
; TITLE OF INVENTION: Polypeptide having immunological
; TITLE OF INVENTION: activity for use as diagnostic reagent and/or vaccine
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KIRBY EADES GALE BAKER
; STREET: Box 3432, Station D
; CITY: Ottawa
; STATE: Ontario
; COUNTRY: Canada
; ZIP: K1M 1H8
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/743,357
; FILING DATE: 21-AUG-1991
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/CA90/00062
; FILING DATE: 23-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Gale, Edwin J.
; REGISTRATION NUMBER: 28,584
; REFERENCE/DOCKET NUMBER: 30924-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 237-6900
; TELEFAX: (613) 237-0045
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1016 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Human Immunodeficiency virus type 1
; STRAIN: BRU
; US-07-743-357-5

Query Match 57.0%; Score 3109; DB 2: Length 1016;
Best Local Similarity 56.3%; Pred. No. 4.6e-257;
Matches 565; Conservative 177; Mismatches 242; Indels 20; Gaps 4;

QY 46 ARE---KTERARETRIQSDRLGTAPRAGGDTIQGATNRLAA-----PQFS 89
Db 13 AREFSEQTRANSPTISSEQTRANSPTREQLQVWGRDNNLSLEAGDRQGTVSFNPQIT 72
QY 90 LMKRPVTVAYTEGQPVVELLDGTGADDSIVAGTELGNNSYSPKIVGIGGIFINTKEYKNVEI 149
Db 73 LMQRPVLTIKIGQLKEALLDGTGADDTVLEEMSLPGRWKPKMIGGIGFIVKVRQYDQILI 132
QY 150 EVLNKKVRATINTGDTPINIFGRNLTALGMSLNPVAKVEPIKMLKPGDKGPKLRWP 209
Db 133 EICGHAKGTVLVGTPTVNIIGNLLTQIGCTNFPISPIETVPVKLPGMDGPKVKWP 192
QY 210 LTKKIEALKETCEKMEKEGLEEAPPNPNYNTPTFAIKKDKNKRMLIDFRELNVQ 269
Db 193 LTEEIKALVEICTEMEKEGKISKIPENYNTPTFAIKKDKTKRWKLVDPFRELNKRQ 252
QY 270 DFTETQLGIPHPAGLAKRRITVLDVGDAYSIPLHEDFRPYTAFTLPSVNNAEFGKRYI 329
Db 253 DFWEVQLGIPHPAGLKKKSVTLVDGDAYFSVPLDEDFRKYTAFTIPSINNETPGIRYQ 312
QY 330 YKVLPGQWKGSIPAIFQHTMRQVLEPFRKANKDVIIIOYMDDLIASDRTDLEHDRVVLQ 389
Db 313 YNLVPGQWKGSIPAIFQSSMTKILEPFRKQNPDIYQYMDLVGSDLEIGQHRTKIEEL 372
QY 390 KELLNGLGFSPPDEFKQDPYPHMGYELWPTKWKQLQIPLQKQIWTVDIQLVGLVNL 449
Db 373 RQHLRWGLTTPDKKHQKEPPLWNGYELHPDKWTVQVPLPEKDSWTVNDIQLVGLN 432
QY 450 WAAQYIPGKTKYKHLRLRGKMTLFEVQWTELAELAEENRIILSOQEGHYOEEKEEL 509
Db 433 WASQIYPGIKYKQLCKLLRGKALTEVPLTEEALELAENREILKEPVHGVYDPSKDL 492
QY 510 EATYQKQOENQWYKIHQEE-KILKVGKYAKVKNTHNGIRLLAQVQKIGKEALVINGR 568
Db 493 TAEIQKQOGQWYIYQEPKLNKTGKYARMGAHTNDVKLTEAVQKITTESIVINGK 552
QY 569 IPKFLPVEREIQWQWNYQVTVIPDWDVFSVTPPLVRLAFNLVGDPIGAETFTDGS 628
```



```

;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Human immunodeficiency virus type 1
; STRAIN: HXB2
; US-07-743-357-22

Query Match          55.0%; Score 2999; DB 2; Length 913;
Best Local Similarity 58.6%; Pred. No. 1e-247;
Matches 535; Conservative 166; Mismatches 208; Indels 4; Gaps 2;

QY 121 IELGNNYSPIKIVGGIGGFINTKEYKNVEIEVLNKKVVRATIMTGDTPINIFGRNLTALGM 180
Db 1 MSLPGRWKPKMIGGIGGFIVRQYDQILIEICGKAIGVILVGPVTPVNIIGNLLTQIGC 60

QY 181 SLNLPVAKVEPIKIMLPGKDGPKLRQWPLTKEKEALKEICEKMEKEGLEEAPTNPY 240
Db 61 TLNFPISPITETVPVKLPGMDGPKVKQWPLTEKIKALVEICTEMEKEGKISKIGPENY 120

QY 241 NPTPTFAIKKDKNKKWMLIDFRELNKVTQDFTFQIGLPHIPAGLAKKRRITVLDVGDYF 300
Db 121 NTPVFAIKKDKSTKRWKLVDFRELNKARTQDFEVQLGIPHPAGLAKKKSVTVLDVGDYF 180

QY 301 SIPLHEDFRPYTAFTLPSVYNNAEPPGKRYIYKVLPGQWKGSPAIFQHTMRQVLEPPRKANK 360
Db 181 SVPLDDEFRKYTAFTIPSIINNTPGIRYOYNVLPQGWKGSPAIFQSSMTKILEPPRKQNP 240

QY 361 DVILIOYMDLILASDRDLEHVRVVLQKELLNGLGFTPDPEKFOKDPPIYHMGYELWP 420
Db 241 DIVIYOYMDLVGSDLEIGQHTKIEELRQHLRWGLTTPDKKHQKEPFFLWMGYELHP 300

QY 421 TKWKLQIKLOPEKIWTVDIOKLGVNLNAAQYPGIKTKLRLIRGKRWLTLEBQWNT 480
Db 301 DKWTQPIVLPKDSWTVDIOKLGVKLNWASQIPIGKVRQLCKLLRTKALTEVIPLT 360

QY 481 ELAEAELENRIILSQEGHYHQBKELEATVQKDQENQWYTKIHQEE-KILKVGKYAK 539
Db 361 EBAELEAENREILKEPVHGVYDPSKDIATAEQOGQGWYQIYOEPEKNLKTKYAR 420

QY 540 VNKTHTNGIRLLAQQVQKIGKALVITWGRIPKPHLPVERIEFQWMDNWTWQVTPDWF 599
Db 421 MGAHTNDVQKTEAVQKITTISIVJWGTPKPKLPDQIKETWETWTWYQATWPEWBF 480

QY 600 VSTPPLVRLAFNLVGPDPGAETFYDGCNRSQSEKAGYVTDGRKDKVKKLEQNTNQ 659
Db 481 VNTPPLVKLWYQLEKPIVGAETFYVDGAAARETKLGAGYVYNGRQKRVVLTDTNQ 540

QY 660 AELEAFAMALTDGPKVNIIVDSQVYVGMGISAQPTSESKIVNQIIEEMIKKEAIVAVY 719
Db 541 TELQATYLLAQDGLGVNIIVTDSQVAGLIQAQDQSESELVQNIIEQLIKKEKVLAWY 600

QY 720 PAHKGIGGQNOVDHLVSQIGRQVLEKLETPAQEBEKEYHSNVKLSHKFGIPNLVAROI 779
Db 601 PAHKGIGGNEQVDKLVLSAGIRKVLFDGIDKAQDEKEYHSNWRAMASDFNLPVYVAKEI 660

QY 780 VNSCAQCOQGAIRHQVNAELGTWQMDCTHLEGKIIIVAVHVASGFIAEYIPQESGRQ 839
Db 661 VASCDKCOLKGAHMGQVDCSPGIWQLDCTHLEGKVLVAVHVASGYIEAEVIPAETGOE 720

QY 840 TALFLKLASRPWTHLHTDNGANTSQEVKVMVNWIGIEQSGFVPYVPSQGVVEAMNH 899
Db 721 TAYFLKLKAGRPVKTTHTDGNSFTGATVRAACWAGIKQFEGIPYVPSQGVVESMKN 780

QY 900 HLKNOISRIREQANTFETILVLAHICMNFKRGGIGDMTPSERLINMITTEQIEQFLQAK 959
Db 781 ELUKTIQVQRDAEKLTAQVQAVFIHNFKRGGIGGYGAGRIYDITATDITQTELQKQ 840

QY 960 NSKLKDFRYFREGDRDLQWKGPELLWKGEGAVLVKVGTDIKIIPRRKAIIRDYGGROE 1019
Db 841 ITKIONFRVYRDSRNSLWKGPAKLWKGAGVAVIQDNSDIKVPVPRKAIIRDYKGOM- 899
```

```

QY 1020 MDSGSHLEGARED 1032
Db 900 --AGDDCVASRQD 910

RESULT 21
US-07-743-357-6
; Sequence 6, Application US/07743357
; Patent No. 5858646
; GENERAL INFORMATION:
; APPLICANT: Kang, Yong C.
; TITLE OF INVENTION: Polypeptide having immunological
; activity for use as diagnostic reagent and/or vaccine
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KIRBY FADES GALE BAKER
; STREET: Box 3432, Station D
; CITY: Ottawa
; STATE: Ontario
; COUNTRY: Canada
; ZIP: K1M 1H8
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/743,357
; FILING DATE: 21-AUG-1991
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/CA90/00062
; FILING DATE: 23-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Gale, Edwin J.
; REGISTRATION NUMBER: 28,584
; REFERENCE/DOCKET NUMBER: 30924-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 237-6900
; TELEFAX: (613) 237-0045
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 913 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; ORGANISM: Human immunodeficiency virus type 1
; STRAIN: MN
; US-07-743-357-6

Query Match          54.2%; Score 2957; DB 2; Length 913;
Best Local Similarity 57.8%; Pred. No. 4.1e-244;
Matches 528; Conservative 166; Mismatches 215; Indels 4; Gaps 2;

QY 121 IELGNNYSPIKIVGGIGGFINTKEYKNVEIEVLNKKVVRATIMTGDTPINIFGRNLTALGM 180
Db 1 MSLPGRWKPKMIGGIGGFIVRQYDQITIGICGKAIGVILVGPVTPVNIIGNLLTQIGC 60

QY 181 SLNLPVAKVEPIKIMLPGKDGPKLRQWPLTKEKEALKEICEKMEKEGLEEAPTNPY 240
Db 61 TLNFPISPITETVPVKLPGMDGPKVKQWPLTEKIKALVEICTEMEKEGKISKIGPENY 120

QY 241 NPTPTFAIKKDKNKKWMLIDFRELNKVTQDFTFQIGLPHIPAGLAKKRRITVLDVGDYF 300
Db 121 NTPVFAIKKDKSTKRWKLVDFRELNKARTQDFEVQLGIPHPAGLAKKKSVTVLDVGDYF 180

QY 301 SIPLHEDFRPYTAFTLPSVYNNAEPPGKRYIYKVLPGQWKGSPAIFQHTMRQVLEPPRKANK 360
Db 181 SVPLDDEFRKYTAFTIPSIINNTPGIRYOYNVLPQGWKGSPAIFQSSMTKILEPPRKQNP 240
```



```
Db      1 PISPIETVPVKLPGMDGPKVQWPLTEEKIKALVEICTEMEKEKISKIGPENPYNPV 60
      245 FAIKKKDKNKRMLDFRELNVKTQDFEIQIGIPHPAGLAKKRRTVLVDGDAFSPIL 304
      61 FAIKKKDSTKWKRLDFRELNVKTQDFEIQIGIPHPAGLAKKRRTVLVDGDAFSPVL 120
      305 HEDFRPYTAFTLPSVNNAPGKRIYKVLPGQWKGSPAIFQHTMRQVLEPFRKANKDVII 364
      121 DEDFRKYTAFTIPSNINNETPGIRYQYNVLPQWKGSPAIFQSMKILEPFRKQNPDI 180
      365 IOYMDILLASDRTLEHDRVVLQKELNGLGFSPTDEKFKDPPYHMGYELWPTKWK 424
      181 YOYMDLLYVGSLEIGQHRTEELRQHLRLGLTPDKKHQKEPFLMWGYELHPDKWT 240
      425 LOKIOLPOKEIWTVDIOKLVGLVNAWAOLYPGIKTKHLCLIRGKMTLVEEYQWTELA 484
      241 VOYIVLPEKDSWTVNDIOKLVGLVNAWAOLYPGIKTKHLCLIRGKMTLVEEYQWTE 300
      485 AELENRIILSQEGBHYQEEKELEATVQKQENQWYKIHQEE-KILKVGKYAKVNT 543
      301 LEAENREILKEPVHGVYDPSKDLJAEIQKOGQGWYQIYQEPFKNLKTGYARMGA 360
      544 HTNGIRLLAQQVVKIGKEALVIGWGRIPKPHLVEREIQWQWMDNYQVTPWDFVSP 603
      361 HTNDVKQLTEAVQKITTESIVWGTGTPKPKLPQKETWETWTEYQWATWPEWEFVNP 420
      604 PLVRLAFNLVGPDPGAEFTFYDGCNRSQSGKAGYVTDRCQKVKVKKLEQTNQOAE 563
      421 PLVKLWYQLEKEPIVGAETFFYVDGGAANRETKLGKAGYVYVNRGRQKVVTLD 480
      664 AFAMALTDGPKVNIIVDSQYVYMGISASOPTESKIVNQIIEEMIKKAIYVAVWPA 723
      481 AIYALQDSGLEVNIVTDSQYALGIIQAPQDSSELYNQIIEQLIKKVKYLANVWPA 540
      724 GIGGNEQVDHLVSGQIRQVLF 745
      541 GIGGNEQVDKLVSGIRKVLFL 562
```

```
RESULT 24
US-09-752-652-1
; Sequence 1, Application US/09752652
; Patent No. 6503123
; GENERAL INFORMATION:
; APPLICANT: Roberts, Grace B.
; APPLICANT: Furfine, Eric S.
; APPLICANT: Porter, David, J.T.
; TITLE OF INVENTION: CONTINUOUS TIME RESOLVED RESONANCE
; TITLE OF INVENTION: ENERGY TRANSFER ASSAY FOR POLYNUCLEIC ACID POLYMERASE
; TITLE OF INVENTION: ACTIVITY
; FILE REFERENCE: PU3761
; CURRENT APPLICATION NUMBER: US/09/752,652
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/167,940
; PRIOR FILING DATE: 1999-11-29
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Human Immunodeficiency virus type 1
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 100
; OTHER INFORMATION: Xaa-Leu or Ile
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 103
; OTHER INFORMATION: Xaa-Lys or Asn
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 106
```

```
; OTHER INFORMATION: Xaa=Val, Ile or Ala
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 106
; OTHER INFORMATION: Xaa=Val or Ile
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 138
; OTHER INFORMATION: Xaa=Glu or Lys
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 181
; OTHER INFORMATION: Xaa=Tyr or Cys
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (0)...(0)
; OTHER INFORMATION: xaa at position 188 can be Tyr or Cys
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (0)...(0)
; OTHER INFORMATION: Xaa at position 236 can be Pro or Leu
US-09-752-652-1

Query Match      34.7%; Score 1889.5; DB 4; Length 560;
Best Local Similarity 60.7%; Pred. No. 4.9e-153;
Matches 340; Conservative 90; Mismatches 129; Indels 1; Gaps 1;

QY      185 PVAKVEPIKMLKPGKDGPKLQWPLTEKIEALKEICEKEKEGLEEAPNPYNTPT 244
      1 PISPIETVPVKLPGMDGPKVQWPLTEEKIKALVEICTEMEKEKISKIGPENPYNPV 60
      245 FAIKKKDKNKRMLDFRELNVKTQDFEIQIGIPHPAGLAKKRRTVLVDGDAFSPIL 304
      61 FAIKKKDSTKWKRLDFRELNVKTQDFEIQIGIPHPAGLAKKRRTVLVDGDAFSPVL 120
      305 HEDFRPYTAFTLPSVNNAPGKRIYKVLPGQWKGSPAIFQHTMRQVLEPFRKANKDVII 364
      121 DEDFRKYTAFTIPSNINNETPGIRYQYNVLPQWKGSPAIFQSMKILEPFRKQNPDI 180
      365 IOYMDILLASDRTLEHDRVVLQKELNGLGFSPTDEKFKDPPYHMGYELWPTKWK 424
      181 XOYMDLLYVGSLEIGQHRTEELRQHLRLGLTPDKKHQKEPFLMWGYELHDXKRT 240
      425 LOKIOLPOKEIWTVDIOKLVGLVNAWAOLYPGIKTKHLCLIRGKMTLVEEYQWTELA 484
      241 VOYIVLPEKDSWTVNDIOKLVGLVNAWAOLYPGIKTKHLCLIRGKMTLVEEYQWTE 300
      485 AELENRIILSQEGBHYQEEKELEATVQKQENQWYKIHQEE-KILKVGKYAKVNT 543
      301 LEAENREILKEPVHGVYDPSKDLJAEIQKOGQGWYQIYQEPFKNLKTGYARMGA 360
      544 HTNGIRLLAQQVVKIGKEALVIGWGRIPKPHLVEREIQWQWMDNYQVTPWDFVSP 603
      361 HTNDVKQLTEAVQKITTESIVWGTGTPKPKLPQKETWETWTEYQWATWPEWEFVNP 420
      604 PLVRLAFNLVGPDPGAEFTFYDGCNRSQSGKAGYVTDRCQKVKVKKLEQTNQOAE 663
      421 PLVKLWYQLEKEPIVGAETFFYVDGGAANRETKLGKAGYVYVNRGRQKVVTLD 480
      664 AFAMALTDGPKVNIIVDSQYVYMGISASOPTESKIVNQIIEEMIKKAIYVAVWPA 723
      481 AIYALQDSGLEVNIVTDSQYALGIIQAPQDSSELYNQIIEQLIKKVKYLANVWPA 540
      724 GIGGNEQVDHLVSGQIRQVLF 743
      541 GIGGNEQVDKLVSGIRKVL 560

RESULT 25
US-09-238-303-9
; Sequence 9, Application US/09238303B
; Patent No. 6284253
; GENERAL INFORMATION:
```

; APPLICANT: Barr, Margaret C.
; TITLE OF INVENTION: No. 6284253el Feline Immunodeficiency Virus Nucleotide Sequence
; FILE REFERENCE: 18617.0059
; CURRENT APPLICATION NUMBER: US/09/238, 303B
; CURRENT FILING DATE: 1999-01-28
; EARLIER APPLICATION NUMBER: US 60/072, 927
; EARLIER FILING DATE: 1998-01-29
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 9
; LENGTH: 1150
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: protein encoded by the pol gene of a recombinant viral
; OTHER INFORMATION: clone constructed from the genomic DNA of a Pallas's cat feline
; OTHER INFORMATION: immunodeficiency virus
US-09-238-303-9

Query Match 33.2%; Score 1810; DB 3; Length 1150;
Best Local Similarity 33.9%; Pred. No. 1.1e-145;
Matches 393; Conservative 222; Mismatches 347; Indels 198; Gaps 18;

QY 8 GPLGKEAPQLPRGPSSAGADTNTPSGSSSGSTGEIYAAREKTERAERETIQGSDRLTA 67
DB 13 GHAGKAAGLAEGESCCPYQTSAAISNSSINNS-----ESATWSINTAFGSSNGVP- 63

QY 68 PRAGGDTIOGATNRGLAAPQFSLWKRPVVTAYIEGQVPEVLLDTGAD-----D 115
DB 64 -----YGLKARDI-----ELIHRPILIIYVNGPIRFLMDTGADITIMNAEDFNILN 110

QY 116 SIVAGIELGNYSKPIVGGIGGFINTKEYKNVEIEVLNKKVRATIMTGDTPI----- 167
DB 111 SIPDGIQ-----TMIGVGGKGRKGRFRRVHLEIRDPNHRQAQCLFGNMCILDDNSLTE 162

QY 168 NIFGRNLTALGMSLNP--VAKVEPIKMLKCGKDGPKLRQWPLFKETKEALKECEKM 225
DB 163 PLLGRONMVRFGAKLVANISNKIPIVKVMKDPKSGPKIKQWPLSKETKEALTEIVYRL 222

QY 226 EKGQLEAEAPTPNPTFAIKKKDKNKRMLIDFRELNVKVTQDTEIQLGPHPAGLA 285
DB 223 EKEGKVRADPNPWTPIFCIKK-SGKWRMLIDFTLNELTEKAEVOLGLPHPAGLQ 281

QY 286 KKRITVLDVGDAFYSIPLHEDPRPYTAFTPLPSVNNAEPKRIYKVLPGWKGSPAIFQ 345
DB 282 ERQQTVDLIDADAYFTIPLDPDAPYTAFTPLPKINNSGGERFVWGLPQGWVLSPLIYQ 341

QY 346 HTMRQVLEPRKANKDVIIOYVDDIILIASDRTDLEHDRVVLQKELLNGLGFSTPDEKF 405
DB 342 STLNNILKPREQHEIDLYQYDDIYIGSDLGKKEHKQIVEELKLLLMWGFETPEDKL 401

QY 406 QKDPYHWMGYELWPTKWKQLQTLQPKETIWTVDIQLVGNVNAQAQYLGKTKHLGR 465
DB 402 QEOPPYKMWGYELPRKWTIQTRELIIPPEPTLNELQKLVGINNSSQIIPGLRIKALTN 461

QY 466 LIRGKMTLTEVQWTELAELAEENRIILSQEGHYOEEKELEANTVQKQDNQWYKI 525
DB 462 MMKGNOALDSKRRTWEAKAEAKLAIEQHTQLGYDPQQLHAKLSIVGPHCIQYQV 521

QY 526 HQS---EKILVKCYAKVKNTHNGIRLLAQVVKIGKEALVIWGRIPKFLPVERIE 582
DB 522 YQKGSFDKILWYKMNQRKKAENTCDIALRIYKIRIESIVRLGKEPIYEIPCSREWE 581

QY 583 QMWDNTYQVWTI----PDMDFVSTPPLVRLAFNLV-GDPIPGAETFTYDGCNRSQKEG 637
DB 582 ---SNLINTPYLACPPQVEYIHAAIMIQRSLSMIKEEPIRGAETWIDGG-RKKQSAR 637

QY 638 AGVTVDRGDKVKKLEOTTNQQAELFAFAMALPDSGPKVINIIVDSQYVWGISASQTESE 697
DB 638 AAYTDTKGWEVWQIE-GSNQRAEVAMALLMARSGGEEMIVTDSQYILNLRKQKPDLM 696

QY 698 SKVNOITIEBMIKKALYIAVWVAHKIGGNQBEVDHL----- 734
DB 697 G-LWQELLEIEIEKVAIFIDWVPGHKIPGNTVDNLQCTMMIISNGIILDKGEEDAGYD 755

QY 735 -----VSQGIQVQL----- 743
DB 756 LLAEQDIHLMPGEVRIVPTGVRMLMPKGMWGMVVGKSSIAKQGLDVLGGVIDEGYRGEIG 815
QY 744 -----FLEK 747
DB 816 VININLQKRSITLKEKQKVAQLIIIPCKHEELAKOGEIELNSERGEKGYGSTGAFASWMNN 875
QY 748 IEPAQEEHEKYHNSVKELSHKFGIPNLVARQIVNSCAQCOQKGEATHGQVNAELGTWQMD 807
DB 876 IEAEINHEKFSHPDFLRTFGLPKQVAEEIRKCPCLCIVQGEQVMGKLKVGPGIWIQID 935
QY 808 CTLEKGIILIVAVHVASGETAEAVIQESGRQATALLKLASRWPTLHLHTDNGANFTSQ 867
DB 936 CTLEKGIILIVAVHVASGETAEAVIQESGRQATALLKLASRWPTLHLHTDNGANFTSQ 995
QY 868 EVKMAVWIGIEQSGFVYPNPQSGVVEAMNHLKQISIRIEQANTTETIVLMAHCMN 927
DB 996 KVEGMTGFLGIKHKGYPGNPQSQALVENTNRMLKEWIKKFRGEVTTLDAALALALYALN 1055
QY 928 FKRRGIGDMTPSERLINMITTEQETQFLOAKNSKLKDRVYFREGDQLWKQPGELLWK 987
DB 1056 FKQGRIGRISPYELLIQOESDRIDYFSKIPANNIKNSWIYYKDRDRDKEMKGTQVEYW 1115
QY 988 GEGAVLVK-VGTDIKIIPRR 1006
DB 1116 GOGAVLIKHPHGYMLIPRR 1135

RESULT 26
US-09-946-239-9
; Sequence 9, Application US/09946239
; Patent No. 6579527
; GENERAL INFORMATION:
; APPLICANT: Barr, Margaret C.
; TITLE OF INVENTION: No. 6579527el Feline Immunodeficiency Virus Nucleotide and
; FILE REFERENCE: 18617.0059
; CURRENT APPLICATION NUMBER: US/09/946, 239
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 09/238, 303, US 60/072, 927
; PRIOR FILING DATE: 1999-01-28, 1998-01-29
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 9
; LENGTH: 1150
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: protein encoded by the pol gene of a recombinant viral
; OTHER INFORMATION: clone constructed from the genomic DNA of a Pallas's cat feline
; OTHER INFORMATION: immunodeficiency virus
US-09-946-239-9

Query Match 33.2%; Score 1810; DB 4; Length 1150;
Best Local Similarity 33.9%; Pred. No. 1.1e-145;
Matches 393; Conservative 222; Mismatches 347; Indels 198; Gaps 18;

QY 8 GPLGKEAPQLPRGPSSAGADTNTPSGSSSGSTGEIYAAREKTERAERETIQGSDRLTA 67
DB 13 GHAGKAAGLAEGESCCPYQTSAAISNSSINNS-----ESATWSINTAFGSSNGVP- 63

QY 68 PRAGGDTIOGATNRGLAAPQFSLWKRPVVTAYIEGQVPEVLLDTGAD-----D 115
DB 64 -----YGLKARDI-----ELIHRPILIIYVNGPIRFLMDTGADITIMNAEDFNILN 110

QY 116 SIVAGIELGNYSKPIVGGIGGFINTKEYKNVEIEVLNKKVRATIMTGDTPI----- 167
DB 111 SIPDGIQ-----TMIGVGGKGRKGRFRRVHLEIRDPNHRQAQCLFGNMCILDDNSLTE 162

QY 168 NIFGRNLTALGMSLNP--VAKVEPIKMLKCGKDGPKLRQWPLFKETKEALKECEKM 225
DB 163 PLLGRONMVRFGAKLVANISNKIPIVKVMKDPKSGPKIKQWPLSKETKEALTEIVYRL 222


```
QY 226 EKEGOLERAPPNTPTFAIKKKDKNKRMLIDFRELNKVYTDFTETQIGLPHPAGLA 285
DB 223 EKEGKVRADPNPNWTFECIKKK-SGKWRMLIDFRLNTEKGAEVQLGELPHPAGLQ 281
QY 286 KKRITVLDVGDAYESIPLHEDFRPYTAFTLPSVNAEPGKRYIYKVLPGQWKGSPAIQ 345
DB 282 ERKQVTLIDADAYETIPLDPYAPYTAFTLPIKNSGPFGRFVWGLPGQWVLSPLIYQ 341
QY 346 HTMROVLEPFRKANKDVIIQYMDLILASDRFDLEHDRVVLQKLKELLNGLGFSPDEKF 405
DB 342 STLNNILKPFREHPEIDLYQYMDIYIGSDLGKKEHQIVBELRKLKLLWNGFETPEDKL 401
QY 406 QKDPYHWMGYELWPTKWLQIOLPQKEIWTNDIQLVGLVNLNAAQLPGIKYKHLCR 465
DB 402 QBPYPYKWMGYELPRKWTIQIKELIPEEPTNLQKLVGLIINSSQIIPGLRIKALTN 461
QY 466 LIRGKMTTEEVQWTELAELNRIILSQBQEGHYOEBKELEATVQKQOENQWTYKI 525
DB 462 MMKGQALDSKRWTTEAKKEAEAKLAIEQHTQLGYVDPQQLHAKLSIVGPHCIGYQV 521
QY 526 HQE---EKLKVGKYAKVKNHTNGIRLLAQVVGKIGKEALYVWGRIPKPHLPVERIWE 582
DB 522 YQKSPDKILWYKGNRKKKAENTCDIALRAIYKIREESIYRLGKEPIYEIYPCREAW 581
QY 583 QWMDNYQWVTWI---PDWDFVSTPPLVRLAPNLV-GDPIPCAETFYTDGSCNROSKEG 637
DB 582 ---SNLINTPYLKACPPQVEYIHAAMIQRSLSMKIEPIRGAETWYIDGG-RKGQSNK 637
QY 638 AGYVDRGDKVYKLEQTNQOAELEAFAMALTDGPKVNIIVDSQYVNGISASQPTSE 697
DB 638 AAYWTDKGVEMQIE-GSNRAEVNALLMALRSGEEMNIIVTDSQYILNLRQKPDLM 696
QY 698 SKIVNQIIEEMKKEAIIYVAVPAHKGGGQNOVDHL----- 734
DB 697 G-LWQEIIEETEKKAIFAIDWPGHGKIPGNTVDNLQCTMMIISNGILDKGEDAGYD 755
QY 735 -----VSGIROVL----- 743
DB 756 LLAQDIHLMGPEVRIVTGVRLMLPKHGWMGVGKSSIAKOGDLVGLGVDEGYRGEIG 815
QY 744 -----FLEK 747
DB 816 VIMINLQKRSITLKEKQKVAQLIIPKHEELKQGEIELNSBERGKGYGSGAFASWMNN 875
QY 748 IEPAQEEHEKYSNVKLSHFGIPNLVARIQVNSCAOQKGEAIHQVNAELGTQWMD 807
DB 876 IEAEINHEKHSDDPEFLTEFGLPKQVAEEIKRKCPLCIVQGEQVMGKLYGPGIWIQID 935
QY 808 CTHLEGGIIVAVHVASGFIKAEVIPSQSGROTALFLKLASRWPITHLHTDNGANFTSQ 867
DB 936 CTHLEGGIILVAVNTESGYIWARIIPOETADMTVYLLQLLSEHHVTELOSDNGPFNNA 995
QY 868 EYKVMVWVWIGIESQSGVPYNPQSGVBSAMNHLKNOISRIREQANTIETIYVMAIHCNM 927
DB 996 KVEGTMGELGKHKHYGIPNQSQUALVENTNRMLEKWKIKRFGVTTILDAALALAYALN 1055
QY 928 FKRRGGIGDMPSERLINNIITEQEIQLQAKNSKLDFRYFYFREGRLQKGPCELLWK 987
DB 1056 FKQGRIGRISPYELLIOQESDRIRDFYSKIPANNIKNSWIYIKDRDRKEMKGPQVYEW 1115
QY 988 GEGAVLVK-VGTDIKIIPR 1006
DB 1116 QGGAVLIXHPEHYMLIPRR 1135
```

RESULT 27

```
US-08-811-682-15
; Sequence 15, Application US/08811682
; Patent No. 6331616
; GENERAL INFORMATION:
; APPLICANT: Tompkins, Wayne A.F.
; APPLICANT: Tompkins, Mary B.
```

```
; APPLICANT: Yang, Joo-Sung
; TITLE OF INVENTION: Feline Immunodeficiency Virus Clone
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell Seltzer Park & Gibson
; STREET: PO Drawer 34009
; CITY: Charlotte
; STATE: No. 6331616th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,682
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Virginia C.
; REGISTRATION NUMBER: 37,092
; REFERENCE/DOCKET NUMBER: 5051-332
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1031 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-811-682-15

Query Match 30.7%; Score 1673; DB 4; Length 1031;
Best Local Similarity 35.3%; Pred. No. 4.9e-134;
Matches 362; Conservative 190; Mismatches 309; Indels 164; Gaps 17;

QY 134 GIGGINTKEYKNVEIENVLNVKVRATIMTGDTPI-----NIGFRNLTALGMSLNL 185
DB 3 GYGGGKRGNTYINVHLEIRDENYKTCIFGNCVLEDSNLIQPLGRD--NMIRFNIRLV 60
QY 186 VA---KVPIKMLKPGDKPKLROWPLTKKIEALKKIECKMEKEGOLLEAPPTNPN 241
DB 61 MAQISDKPIVAVKVKDPNPKGQIKQWPLTNKIEALTEIVERLEGEKVKRADPNPN 120
QY 242 TPTFAIKKKDKNKRMLIDFRELNKVYTDFTETQIGLPHPAGLAKRRITVLDVGDAYFS 301
DB 121 TPTFAIKKK-SGKWRMLIDFRELNKLTKEGAEVQLGLPHPAGLKKKQVTVLDIGDAYFT 179
QY 302 IPLHEDFRPYTAFTLPSVNAEPGKRYIYKVLPGQWKGSPAIQHTMROVLEPFRKANKD 361
DB 180 IPLDPDYAPYTAFTLPRKNNAGPGRYVWCSLPQGWVLSPLIYQSLTDNIIOFFIRQNPE 239
QY 362 VTIIOYMDLILASDRFDLEHDRVVLQKLKELLNGLGFSTPDEKFOKDPYHWMGYELWPT 421
DB 240 LDIQYMDDIYIGSNLSKKEKHEKVEELRKLKLLWNGFETPEDKLOEPEYKWMGYELHPL 299
QY 422 KWKLOKIQLPQKEIWTNDIQLVGLVNLNAAQLPGIKYKHLCRIRGKMTITEEVQWTE 481
DB 300 TWSIQOQKQLNEIPERPTNLQELQKAGINNASQITPDLSTIKELTNMNRGDKLDSIREWTV 359
QY 482 LAEAELENRIILSQBQEGHYOEBKELEATVQKQOENQWTYKIHQE--EKLKVGKYAK 539
DB 360 EAKREVQKAEIEMQAQLNYDHPRELYAKLSLVGPHQICQVYVHKNEPCTLWYGKMNR 419
QY 540 VKNHTNGIRLLAQVVGKIGKEALYVWGRIPKPHLPVERIWEQWMDNYQWVTWI---P 595
DB 420 QKKKAENTCDIALRACYKIREESIYRLGKEPIYEIYPTSEAW---SNLINSPLYKAPPP 476
QY 596 DMDFVSTPPLVRLAPNLVGD-PIPGAETFYTDGSCNROSKEGKAGYVTDGRKDKVKKLEQ 654
DB 477 EVEYIHAAVNIKRALSMTIKDVPIPEAETWYIDGG-RKLGAAGAAYWTDTGKQWMELE- 534
```


Db 1 QWPLTEKIKALVEICTEMEKEGKISKIGPENPNTPVFAIKKDKSTKWRKLVDFRELNK 60
Qy 267 VTQDTEITQIGIPHPAGLAKKRRTITVLDVGDAVESIPLHEDFRPYTAFTPLSPVNAAPGK 326
Db 61 RTQDFWEVQIGIPHPAGLAKKRRTITVLDVGDAVESVPLDEDFRYTAFTPLSPVNAAPGK 120
Qy 327 RYIKVLPQGWKSPAIHQHTMRQVLEPFRRKANKDVIIIOYMDLILASDRDLEHDRV 386
Db 121 RYQNVLPQGWKSPAIHQHTMRQVLEPFRRKANKDVIIIOYMDLILASDRDLEHDRV 180
Qy 387 LQKELLNGLGFSPTDEKFKQDPYHMGVELWTKWKLQIKLQIPQKEITVNDIQKLVG 446
Db 181 EELQHLRLMGLTTPDKKHQKPEPFLMWGVELHDPKWTQVIVLPEKDSWTVNDIQKLVG 240
Qy 447 VLNAAOLYPGIKTKHLRLIRGKWTLEEVQVMTLAEAELEENRIILSQBQEGHYOEE 506
Db 241 KLNWASQIYPGIKTKHLRLIRGKWTLEEVQVMTLAEAELEENRIILSQBQEGHYOEE 300
Qy 507 KELEATVOK 515
Db 301 KDLIAEQK 309

RESULT 32

US-08-987-867A-6
; Sequence 6, Application US/08987867A
; Patent No. 6063384
; GENERAL INFORMATION:
; APPLICANT: C. Morrow et al.
; TITLE OF INVENTION: ENCAPSULATED RECOMBINANT VIRAL
; TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 STATE STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/987,867A
; FILING DATE: 09-DEC-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/087,009
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: UAP-004CPDV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 314 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-987-867A-6

Query Match 19.6%; Score 1068; DB 3; Length 314;
Best Local Similarity 63.4%; Pred. No. 4.4e-83;
Matches 196; Conservative 41; Mismatches 72; Indels 0; Gaps 0;
Qy 207 QWPLTEKIKALVEICTEMEKEGKISKIGPENPNTPVFAIKKDKKWRKLVDFRELNK 266

Db 1 QWPLTEKIKALVEICTEMEKEGKISKIGPENPNTPVFAIKKDKSTKWRKLVDFRELNK 60
Qy 267 VTQDTEITQIGIPHPAGLAKKRRTITVLDVGDAVESIPLHEDFRPYTAFTPLSPVNAAPGK 326
Db 61 RTQDFWEVQIGIPHPAGLAKKRRTITVLDVGDAVESVPLDEDFRYTAFTPLSPVNAAPGK 120
Qy 327 RYIKVLPQGWKSPAIHQHTMRQVLEPFRRKANKDVIIIOYMDLILASDRDLEHDRV 386
Db 121 RYQNVLPQGWKSPAIHQHTMRQVLEPFRRKANKDVIIIOYMDLILASDRDLEHDRV 180
Qy 387 LQKELLNGLGFSPTDEKFKQDPYHMGVELWTKWKLQIKLQIPQKEITVNDIQKLVG 446
Db 181 EELQHLRLMGLTTPDKKHQKPEPFLMWGVELHDPKWTQVIVLPEKDSWTVNDIQKLVG 240
Qy 447 VLNAAOLYPGIKTKHLRLIRGKWTLEEVQVMTLAEAELEENRIILSQBQEGHYOEE 506
Db 241 KLNWASQIYPGIKTKHLRLIRGKWTLEEVQVMTLAEAELEENRIILSQBQEGHYOEE 300
Qy 507 KELEATVOK 515
Db 301 KDLIAEQK 309

RESULT 33

US-08-679-493A-72
; Sequence 72, Application US/08679493A
; Patent No. 6303295
; GENERAL INFORMATION:
; APPLICANT: Taylor, Ethan W.
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
; FILE REFERENCE: 55-95
; CURRENT APPLICATION NUMBER: US/08/679,493A
; CURRENT FILING DATE: 1996-07-12
; PRIOR APPLICATION NUMBER: 60/001203
; PRIOR FILING DATE: 1995-07-14
; PRIOR APPLICATION NUMBER: 60/003,112
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 72
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(327)
; OTHER INFORMATION: X is selenocysteine.
US-08-679-493A-72

Query Match 19.4%; Score 1055; DB 4; Length 327;
Best Local Similarity 59.5%; Pred. No. 6.1e-82;
Matches 198; Conservative 44; Mismatches 79; Indels 12; Gaps 2;
Qy 185 PVAKEPIKIMLPGKDGPKLROMPLTKERIEALKEICEKMEKEGQLEEEAPPNTPT 244
Db 1 PISPIETVPVKLPGMGDPKVKQWLPTEEKIKALVEICTEMEKEGKISKIGPENYTPV 60
Qy 245 FAIKKDKKWRMLIDFRELNKVTQDTEITQIGIPHPAGLAKKRRTITVLDVGDAFSIPL 304
Db 61 FAIKKDKSTKWRKLVDFRELNKRTQDFWEVQIGIPHPAGLAKKRRTITVLDVGDAFSIPL 120
Qy 305 HEDFRPYTAFTPLSPVNAAPGKRYIKVLPQGWKSPAIHQHTMRQVLEPFRRKANKDVII 364
Db 121 DEDFRPYTAFTPLSPVNAAPGKRYIKVLPQGWKSPAIHQHTMRQVLEPFRRKANKDVII 180
Qy 365 IQYMDLILASDRDLEHDRVVQVLEPFRRKANKDVIIIOYMDLILASDRDLEHDRV 418
Db 181 LSTHGXFVCRIXLEIGHQHTKIE-----ELRQHLLRWGLTTPDKKHQKPEPFLMWGVEL 234
Qy 419 WPTKWLQIKLQIPQKEITVNDIQKLVGNWAAQLYPGIKTKHLRLIRGKWTLEEVQV 478
Db 235 HPDKWTQVIVLPEKDSWTVNDIQKLVGNWAAQLYPGIKTKHLRLIRGKWTLEEVQV 294

```
QY 479 WTELAEELEENRIILSOEQEGHYOEKELEA 511
  ||||| ||||| ||||| ||||| ||||| |||||
Db 295 LTERAELEAENREILKEPVHGYIDPSKDLIA 327
  ||||| ||||| ||||| ||||| ||||| |||||

RESULT 34
US-08-679-493A-71
; Sequence 71, Application US/08679493A
; Patent No. 6303295
; GENERAL INFORMATION:
; APPLICANT: Taylor, Ethan W.
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
; FILE REFERENCE: 55-95
; CURRENT APPLICATION NUMBER: US/08/679,493A
; CURRENT FILING DATE: 1996-07-12
; PRIOR FILING DATE: 1995-07-14
; PRIOR APPLICATION NUMBER: 60/001203
; PRIOR FILING DATE: 1995-07-14
; PRIOR APPLICATION NUMBER: 60/003,112
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 71
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; NAME/KEY: VARIANT
; LOCATION: (1)..(327)
; OTHER INFORMATION: X is selenocysteine.
US-08-679-493A-71

Query Match 19.2%; Score 1048; DB 4; Length 327;
Best Local Similarity 59.0%; Pred. No. 2.4e-81;
Matches 193; Conservative 46; Mismatches 86; Indels 0; Gaps 0;

QY 185 PVAKVEPIKIMLKPGDKPKLRQWPLTKKEIKALKEICERKEGQLEAEAPPTNPYPT 244
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 PISPETVPVKLKPGDKPKVQWPLTEEEKIKALVEICTEMEKEGKISKIGPENYPNPV 60
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 245 FAIKKKNKRWMLIDFRELNVQDFTETIQLGIPHPAGLAKKRITVLDVGDAYFSPL 304
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 FAIKKDKSTWRKVLDFRELNKTQDFEWOLGIPHPAGLAKKRISNSTGCGXCIFSSL 120
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 305 HEDFRPYTAFTLPSVNAEPCKRYIYKVLPGWKGSPAIFQHTMRQVLEPPRKANKOVII 364
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 DEDFRKYTAFTIPINNTPGIRYQYNVLPQGWKSPAIFQHTMRQVLEPPRKANKOVII 180
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 365 IQYMDLILASDRTDLEHDRVVLQKELLNGLGFSTPDEKFKDPPYHMGYELMPTKWK 424
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 YOYMDLVGSDLEIGQHRTKIEELRQHLRWGLTTPDKKHQKEPPFLMGMGYELHPDKWT 240
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 425 LQKIQPKETWTVNDIQKLVGLNWAQAQLYPGIKTKHLCKLRGKWTLTTEVQWTELAE 484
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 VQIYVLPKDSWTVDIQKLVGLNWAQAQLYPGIKTKHLCKLRGKWTLTTEVQWTELAE 300
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 485 AELEENRIILSOEQEGHYOEKELEA 511
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 LEAENREILKEPVHGYIDPSKDLIA 327
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 35
5320958-3
; Patent No. 5320958
; APPLICANT: INOUE, SUNIKO; HSU, MEI-YIN; EAGLE, SUSAN;
; INOUE, MASAYORI
; TITLE OF INVENTION: ISOLATED BACTERIAL REVERSE TRANSCRIPTASE
; NUMBER OF SEQUENCES: 24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/315,316
; FILING DATE: 24-FEB-1989
; SEQ ID NO:3
; LENGTH: 261

Query Match 17.0%; Score 924.5; DB 4; Length 290;
Best Local Similarity 57.1%; Pred. No. 7.5e-71;
Matches 165; Conservative 60; Mismatches 61; Indels 3; Gaps 1;

QY 744 FLEKIEPAQEHEEKYHSHVNSKELSHKFGIPNLVARQIVNSCAQCKQGEAIGHQVNAELGT 803
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3 FLGDIDKAQEHEEKYHSHNRAMASDFNLPPVVAKEIVASCQCKQLKGEAMHGQVDCSPGI 62
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 804 WQMDCTHLEGKIIIVAVHVASGFTAEAVIPOESGQWALFLKLLKASRPITHLHNDGAN 863
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 63 WQMDCTHLEGKIIIVAVHVASGFTAEAVIPOESGQWALFLKLLKASRPITHLHNDGAN 122
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 864 FTSEVKNVAVWIGIEQSGFVPYPOSGVYEAAMHHLKNOISRRECAANTTETVLMAI 923
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 123 FTSITVKAACWAGIKQEGFIPYPOSGVYEAAMHHLKNOISRRECAANTTETVLMAI 182
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 924 HCMNFKRRGIGDMTPSPERLINMITTEQEIQFLOAKNSKLDKFRVYFREGDQLKGPGE 983
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```
US-07-648-796A-6
; Sequence 6, Application US/07648796A
; Patent No. 5310876
; GENERAL INFORMATION:
; APPLICANT: Dr. Hubert Bayer
; TITLE OF INVENTION: Expression of HIV1 and HIV2
; TITLE OF INVENTION: Polypeptides and their use
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/648,796A
; FILING DATE: 19910125
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 40 02 636.1
; FILING DATE: 30-JANUARY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5310876man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 784
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acid residues
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-648-796A-6

Query Match 15.2%; Score 827.5; DB 1; Length 376;
Best Local Similarity 54.2%; Pred. No. 2.4e-62;
Matches 156; Conservative 56; Mismatches 59; Indels 17; Gaps 4;

QY 740 ROVLFE-----KIPAEHEKYSNVKELS--HKFGIPNLVARQIVNSCAQCOQKGEA 792
DB 86 KQVRYLEANISKLSQAQIQEK--NMVELQKLSNWDPP-----LESCKDKQLKGEA 135

QY 793 IHGOVNAELGTWQMDCTHLEGLIIVAVHVASGFIEAEVPOESGROTALFLKLASRWP 852
DB 136 MHGQVDCSPGIWQDCTHLEGLIIVAVHVASGFIEAEVPAETGOETAYFIKLGRWP 195

QY 853 IPHLHTDNGANTSOEVKVMVNWIGESQGVYPNQSQGVVYAEAMNHLKNOISIREQA 912
DB 196 KVIHTDNGSNFTSTVRAACWAGIKQEFGIPYNPQSQGVVYAEAMNHLKNOISIREQA 255

QY 913 NTIETVLMIAHMCNFKRGGDMTPSERLINMITTEQEIQLQAKNSKLKDFRYFRE 972
DB 256 EHLKTAVQMAVFIHFKRGGIGGYSAGERIWDIATDIQTKELQKIIKIQNFERYRD 315

QY 973 GRDQLWKGPELLWGEAVLVKGTDIKIIPRRKAKIIRDYGGREQM 1020
DB 316 SRDPLWKGPAKLLWKGAVVQDNSEIKVPRRKAIRDYGKQSDL 363

RESULT 40
5252477-3
; Patent No. 5252477
; APPLICANT: OROSZIAN, STEPHEN; COPELAND, TERRY D.
; TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS SPECIFIC

; PROTEOLYTIC ENZYME AND A METHOD FOR ITS SYNTHESIS AND RENATURATION
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/57,183
; FILING DATE: 01-JUN-1987
; SEQ ID NO: 3
; LENGTH: 275
5252477-3

Query Match 15.1%; Score 821; DB 6; Length 275;
Best Local Similarity 56.2%; Pred. No. 5e-62;
Matches 150; Conservative 43; Mismatches 72; Indels 2; Gaps 1;

QY 55 RETIQSDRGLTAPRAGGDTIQCATNRLGAAPQFSLWKRVPVVTAYIEGQVPEVLLDTGAD 114
DB 9 RRELQVWGRDNNSPSEAGADROGTVSFNF--PQITLWQRPLVTIKIGGLKEALLDTGAD 66

QY 115 DSIVAGIELGNYSKPIGVGIGGFINTKYNKVEIEVLNKKVRAITMTGDTPINIFGRNI 174
DB 67 DTVLEMSLPGRWKPKMIGGIGGFIVROYDQILIEICGHKAIGTVLVGPTPVNIIGRNL 126

QY 175 LTAGMSLNLPAKVEPIKIMLKPGKDGPKLQWPLTKKIEALKEICEKMEKEGLEEA 234
DB 127 LTQIGCTLNFPIETVPEVKLPGMDGPKVQWPLTELIKALVEICTEMEKEGKISKI 186

QY 235 PPTNPYNTPTFAIKKKDKNKRMLIDFRELNKVTDFTETIQLGIPHPAGLAKKRRITVLD 294
DB 187 GPENYNTPTVFAIKKKDKSTKWRKLVDFRELNRKTDQDFWEVQLGIPHPAGLAKKRSVTVD 246

QY 295 VGDVFSIPLHEDFRPYTAFTLPSVNN 321
DB 247 VGDVFSVPLDEDFRKYTAFTIPSINN 273

RESULT 41
US-07-648-796A-2
; Sequence 2, Application US/07648796A
; Patent No. 5310876
; GENERAL INFORMATION:
; APPLICANT: Dr. Hubert Bayer
; TITLE OF INVENTION: Expression of HIV1 and HIV2
; TITLE OF INVENTION: Polypeptides and their use
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/648,796A
; FILING DATE: 19910125
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 40 02 636.1
; FILING DATE: 30-JANUARY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5310876man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 784
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acid residues
```

;
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-648-796A-2

Query Match 14.7%; Score 804; DB 1; Length 247;
Best Local Similarity 60.9%; Pred. No. 1.2e-60;
Matches 142; Conservative 45; Mismatches 46; Indels 0; Gaps 0;

QY 783 CAQCOQGEATHGQVNAELGTWQMDCHLEKIIIVAVHVHVASGFIEAEVIPAQESGRTAL 842
DQ 8 CDKCOLGAEAMHGQVDCSPGWLQDCHLEKIIIVAVHVHVASGFIEAEVIPAQESGRTAY 67
QY 843 FLLKSLRWPIITLHTDNGANFTSQEVKVAWMIIGIEQSGFVYPNQSOGVVEAMNHLK 902
DQ 68 FILKLAGRPVKVLIHTDNGSNFTSTTVKAACWAGIKQEGIPYNPQSGVWESMKNELK 127
QY 903 NQISRIREQANTITVLMIAHNCNFKRGGIGDMTPSERLINMITTEQIEQIFQAKNSK 962
DQ 128 KIIGQVRDQAEHLKTAQVMAVFIHFKRKGIGGYSAGERIVDIATDIOTKELQKQIK 187
QY 963 LKDFRVYFRGRDOLWKGPCELLWKGAVLVKVGTDIKIIPRKAKIIRDYG 1015
DQ 188 IQNFRVYRSDRLPWKGPALKLWKGEVAVIQDNSEIKVYVPRKAKIIRDYG 240

RESULT 42

US-08-679-493A-73
; Sequence 73, Application US/08679493A
; Patent No. 6303295

GENERAL INFORMATION:

; APPLICANT: Taylor, Ethan W.
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS

; FILE REFERENCE: 55-95
; CURRENT APPLICATION NUMBER: US/08/679,493A

; CURRENT FILING DATE: 1996-07-12

; PRIOR APPLICATION NUMBER: 60/001203

; PRIOR FILING DATE: 1995-07-14

; PRIOR APPLICATION NUMBER: 60/003,112

; PRIOR FILING DATE: 1995-09-01

; NUMBER OF SEQ ID NOS: 216

; SOFTWARE: PatentIn ver. 2.0

; SEQ ID NO 73

; LENGTH: 237

; TYPE: PRT

; ORGANISM: Human immunodeficiency virus type 1

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: (199)

; OTHER INFORMATION: X is Q or Y.

; NAME/KEY: VARIANT

; LOCATION: (1)..(237)

; OTHER INFORMATION: X at positions 186, 192, 237 and 233 is

; OTHER INFORMATION: selenocysteine.

US-08-679-493A-73

Query Match 12.7%; Score 691; DB 4; Length 237;
Best Local Similarity 70.5%; Pred. No. 5.3e-51;
Matches 122; Conservative 23; Mismatches 28; Indels 0; Gaps 0;

QY 185 PVAKVEPIKMLPGKDGPKLRWPLTKKIEALKEICEKMEKEGQLEAPPTNPNYPT 244

DQ 1 PISPIETVPVKLPGMDGPKVQWPLTEERIKALVEICTEMEKEGKISKIGPENPNTPV 60

QY 245 FAIKKDKNKRMLIDPRELNKYTDPTETQLGPHIPAGLAKRRITVLDVGDAYSIPL 304

DQ 61 FAIKKDKSTKRWKLVDPRELNKYTDQDFEWQLGPHIPAGLAKRRITVLDVGDAYSVPL 120

QY 305 HEDFRPYTAFTLPSVNAEPGKRYIYKVLPGQWKGSPAIQHTMROVLEPFRK 357

DQ 121 DEDFRKYTAFTIPSINNETGIRYQYNLPGQWKGSPAIQSSMTKILEPFRK 173

RESULT 43

US-08-259-451-11

; Sequence 11, Application US/08259451

; Patent No. 6406841

GENERAL INFORMATION:

; APPLICANT: Lee, Helen H.

; APPLICANT: Swanson, Priscilla A.

; APPLICANT: Idler, Kenneth B.

; APPLICANT: Rosenblatt, Joseph D.

; APPLICANT: Chen, Irvin S. Y.

; APPLICANT: Golde, David W.

; APPLICANT: Robertson, Eugene F.

; APPLICANT: Stephens, John E.

; APPLICANT: Chan, Emerson W.

; APPLICANT: Buytendorp, Mark H.

; APPLICANT: Johnson, Joan E.

; APPLICANT: Motley, Cheryl T.

; APPLICANT: Peterson, Bryan

; APPLICANT: Edwards, Michelle

; APPLICANT: Guidinger, Peggy

; APPLICANT: Tate, Cynthia

; TITLE OF INVENTION: HTLV-II/III Compositions

; TITLE OF INVENTION: and Assays for Detecting HTLV Infection

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Abbott Laboratories

; STREET: One Abbott Park Road

; CITY: Abbott Park

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60064

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb

; MEDIUM TYPE: storage

; COMPUTER: IBM

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WordPerfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/259,451

; FILING DATE: 20-JUN-1994

; CLASSIFICATION: 436

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/086,415

; FILING DATE: 01-JUL-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Daniel W. Collins

; REGISTRATION NUMBER: 31,912

; REFERENCE/DOCKET NUMBER: 5381.US.PI

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (708) 937-6365

; TELEFAX: (708) 938-2623

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 917 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: unknown

US-08-259-451-11

Query Match 10.8%; Score 588.5; DB 4; Length 917;
Best Local Similarity 25.1%; Pred. No. 3.2e-41;
Matches 233; Conservative 144; Mismatches 364; Indels 188; Gaps 39;

QY 197 KPCKDGPKLRQWPLT--KEKIEALKEICEKMEKEGQLEAPPTNPNYNTPTFAIKKDKNK 254

DB 37 RPTPTPTAGPISPKPKRLQALNDLVSKALEAGHIE--PYSGGNNPVPVKKN-GK 93

QY 255 WRMLIDPRELNKYT-----QDTEIQLGPHIPAGLAKRRITVLDVGDAYSIPL 304

DB 94 WRHIDLRLATNATTTITSPSPGPPDLTSLPTALPH-----LQTDLTDAFTQIPL 144

QY 305 HEDFRPYTAFTLPSVNAEPGKRYIYKVLPGQWKGSPAIQHTMROVLEPFRKANKDVII 364


```
Db 145 PROQFYFAFTIPQPCNYGEGTRYAWTVLPQGEKNSPTLFEQQAAAVLPMRKMFTSTI 204
QY 365 IQYMDILLASDRDLEHVRVVLQKLKELLNGLFSTPDEKFKDP-----PYHW 413
Db 205 VOYMDILLASP-TNKLQOLSLTQALTTHGLPISEKQTQTPQIRFLGOVISPNH- 262
QY 414 MGVELMPTKWKLOKIQLPQKEIWTVDIOKLVGLVNWAAQLYPGIKTKHLCLRI----- 467
Db 263 ITVESTPA-----IPIKSQWTLTELQVILGEIQWVSKGTP-ILRKHLQSLYSALHY 313
QY 468 ---RGRWTLT-ERVQWTELAARELEN---RI-----ILSQEGHYQEKELEAT 512
Db 314 ROPRACITLTPQOLHALHAIQALQHNCRGRLDPTLLGLLSLSLSTG-----TTSV 365
QY 513 VOKDQEN---QWTYKIQHEKILKVGKAKVKNTHNGIRLLAQQVOKIGKEALVIWGR 569
Db 366 IFQPKQWPLAWLHTPHPTSLCPWG-----HLLACTILTIDKYTLQHYGLL 412
QY 570 -PKFHLPVREIWEQWMDNY-----WQ-VTWIPDMDVFST 602
Db 413 CQSFHNMKSQALCDPLRNSPHSPVSGILIHMGREHNLGSGPSGPKTLLHLP--TLQ 470
QY 603 PPLVRLAFNLVGDPIGATFTYDGCNRSQKEGKAGYV---TDGKDKVKKLEQTN- 657
Db 471 PLLRPIFTLSPVLDTPCLFSDGS-----PQKAAVYLDQITLQODITPLPPHETNS 524
QY 658 -QOAELEA--FAMALTDSPGVNIIYDQYV-----MGISASQPTSESKIVN 702
Db 525 AQKCELLALIYGLRAAKPWSLNFILDSKYLKYLHSLAIGAPLGTSAHQTLQA----- 578
QY 703 QITEEMIKKAIYVAVWPAHKIGG-----NOEVDHLVSQGINQVLFLEKIEPAQEHEK 757
Db 579 -ALPPLLOGKTVLHVHRSHTNLPDPISFTNEYTDSL-----VAPLVPLTP-QGLHGL 630
QY 758 YHSNWKELSHKGIPNLVARQIVNSCAQQOQGEATH---GQVNAEL---GTWQMDCTHL 811
Db 631 THCNQALY-SFGATPKAKSLVQTCQIINSQHMPQGHIRGLLPNHNHMQGDVTHY 689
QY 812 EKKIIVAVHV---ASGFIEAEVPOESGROT-ALFLKLKASRWPIHLHTDNGANETS 866
Db 690 KYKKYKYCHLVWVDVFGAVSVCKKETSCTISAFLOAISLLGKPLHINDNGAPFLS 749
QY 867 QEVKVMWVIGIEQSGVPYNPQSGVVEAMNHLKNOISR-IREQAN-TIETIVLMAIH 924
Db 750 QEFOECTSYHIKHSHPHPYNTSSGLVERTNGIKNLLNKVLLDCPNLPDINAINKALW 809
QY 925 CMNFRRGIGDMTPSERLINMITTEQEIQFOAKNSKLDFRYVREG-----RDQLWKG 980
Db 810 TLN-----QLNVNPNPQSGKTRWQIHHSPPPLPIPEASTPPKPSKWFFYKPLGLTNQWKG 864
QY 981 PGLLWKGEGAVLVKVGTDIKIIPRRKAK 1009
Db 865 PLSLQEAAGAALLSIDGPRWIPWRFLK 893
```

RESULT 44

US-09-075-272-4

Sequence 4, Application US/09075272

Patent No. 6136598

GENERAL INFORMATION:

APPLICANT: MILLER, A. DUSTY

APPLICANT: WOLGAMOT, GREG

APPLICANT: BONHAM, LYNN

TITLE OF INVENTION: MUS DUNNI ENDOGENOUS RETROVIRAL

TITLE OF INVENTION: PACKAGING CELL LINES

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESS: TOWNSEND and TOWNSEND and CREW LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: U.S.A.

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/075,272

FILING DATE: 08-MAY-1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/046,140

FILING DATE: 09-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Poor, Brian W.

REGISTRATION NUMBER: 32,928

REFERENCE/DOCKET NUMBER: 14538A-003710

TELEPHONE: (206) 467-9600

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1203 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-075-272-4

Query Match

Best Local Similarity 10.6%; Score 576.5; DB 3; Length 1203;

Matches 258; Conservative 190; Mismatches 413; Indels 313; Gaps 45;

```
QY 94 PVVTAYIEGQPEVLELDTGADDSIVAGIELGNYSK--IVGGIGGFI---NTKEYKNVE 148
Db 11 PVTLSVEGTPVNFLLDTGAHSLVTS-PLGLGSKRTIVGVATGSKLIPWTKRALQID 69
QY 149 IEVLNKKVRATIMTGDTPINIFGNILTAL----- 178
Db 70 ---KNMVTSHFLVPECPAPLLGRDLLTKLKAQVQFTSEGPQVSWGKAPLACLVLSTEE 126
QY 179 -----GMSLNLPVAKYEPKIMLPGKDGPKLQWPL 210
Db 127 YRLHEBPQGAAPLDWVTAPPNNVABEQAGMGL---AKQPPVVVVELKADATPISVRYPM 183
QY 211 TREKIEALKEICEKMEKQLEAAPTNPNTPTFAIKKKKNKWRMLIDFRELNKVTD 270
Db 184 SKEAKGICPHIRRLDQGLIYAC--QSPWNPPLPVKPGTNDYRPVQDLREVNKRVL 241
QY 271 FTEIQIGIPHPAGL-----AKRRITVLDVGDAYFISIPLHEDFRPYTATFLPSVNNAP 325
Db 242 ---IHTPTVPNPNLLSSLPPEPTWTYVLDLKDFAFFCLRLHPSQLLFAFEWRDPEGGQT 298
QY 326 KRYIYKVLPGQWKGSPAIQHTMRQVLEPFRANKDVIIIOYMDILLIASDRTDLEHVR 385
Db 299 -QLTWTRLPQGFKNSTPLEDEALHRLAPFRAQNPQLTLQYVDDLLIAAASKELCQQT 357
QY 386 VLQKELLNGLGFTPDEKFO-KDPPYHWMGYEL-----WPTKWKLQKLTQLPQKEITV 439
Db 358 ERLTELGN-LGYRYSAKKAQICQTEVILGTTLRCGKRWLFEARKKTVWMLPPPT-TP 415
QY 440 DIQKLVGLNWAQAQLYPGIKTKHLCLIRGKMTLTEE---VOWTE----- 481
Db 416 QVREFLGTAGFCRLWIPGFAT-----LAAPLYPLTREGIPFEWKEHQFAEIKSSLMT 470
QY 482 ---LAAEAELENRIILSQEGHYQEKELEATVQKDOENQWTKIHOEEKILKVGKVA 538
Db 471 APALAPDLTKSFVL-----YVDERAGIARGVLTQALGPKWRPVAYLSK--KLDPPVA 520
QY 539 KVKNTHTNGIRLLAQQVQ-----KIGKEALVTWGRIPKPHLPVEREIEWQMDN----Y 589
Db 521 SCWPTCLKAIAAVALLIKDADKLTWGOQVTVV---APHALESIVRQPPRWTNARMTHY 577
QY 590 QVTWIPDMDVFSTPPLVRLAF-----NLVGDPIPGAETFF 623
```

```
Db 578 QSLLLNDVTFAPPAILNPATLLPLTNDSDVPVHRCADILAEIGETRKDLTDQWPWGPAPSW 637
Qy 624 YTDGSCNRQSKGAGVYTDGDKV---KKLEOTTNQAELEAFAMALDTS-GPKVNII 679
Db 638 YTDGSSFLIEGKRAGAAVDGKVIWASALPEGSTSAQKAELIALTOALREAEKIINIY 697
Qy 680 VDSQYVMG-----ISASQPTSESKIYVNOIIEEMIKKEAIYVAVWPAH-K 723
Db 698 TDSRYAFATAHIHGAIYRQGLLTSAGKDINKREIL-ALLEAIHAPKKVAILHCPGHQK 756
Qy 724 G-----IGNGNEVDHL---VSQG-----TRQV-----LPLEK 747
Db 757 GEDLVAKGNRMADSVAKOVAQAGAMILTEKGNPKSPEDENYDIKELEFWTSDPLPYFPEGK 816
Qy 748 IEPAQOEHEHYSHNVKLSH-----KFGIPNL--VARQIVNSCAOC----- 786
Db 817 IDLTPEGIEFVKGLHGFTHLGEKMMRLIKKSRYQVFNLSKVAOKIINSCKACAFNAT 876
Qy 787 -----QKGE--AIHQGVNAELGTWQDCTHLE-----KIIIVAVHVSFGFBAEV 831
Db 877 KTYKEPGKRGDRPGVY-----WEVDTEVKPMYGNKYLLVFDVTFSGWVEAFP 927
Qy 832 IPQSGRQATALL-KLASRWPIH-LHTDNGANFTSQEYKVMVNWIGIEQSFGVYVNPQ 889
Db 928 TKTETAQIVAKKIFEEILPRYGVPKVIGSDNGPAFVAQVSGGLATQGLIDMKLHCAVRPQ 987
Qy 890 SOGVVEAMNHLKNOISRIEQANTIETIVLMAI---HCNFKRRGGI----- 934
Db 988 SSGOVERMNTLKETLKLAMEYGGKDWALLPLAFRANTPGRFGLTFEVLVGGPPP 1047
Qy 935 -----GDMTPSERLINM-----ITTEQEIOTLOAKNSKLKDFRVYVREG 973
Db 1048 LIKDGTLVPSGSLVPSLLIHLKALKVIRTQIWDQLKATYTPGTTAVPHEFQV----- 1102
Qy 974 RDQLW-----KPGCELLWKGEVAVLVKVGVDIKI 1002
Db 1103 GDQVLVRRHRTGSLERWKGPYLLVLLTPTAVKV 1136

RESULT 45
US-09-309-572-13
; Sequence 13, Application US/09309572
; Patent No. 6440730
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral hybrid vectors pseudotyped with LCMV
; FILE REFERENCE: P50489
; CURRENT APPLICATION NUMBER: US/09/309,572
; CURRENT FILING DATE: 1999-05-11
; EARLIER APPLICATION NUMBER: DE 198 56 463
; EARLIER FILING DATE: 1998-11-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1737
; TYPE: PRN
; ORGANISM: Moloney murine leukemia virus
; FEATURE:
; OTHER INFORMATION: gag-pol protein
US-09-309-572-13

Query Match 10.2%; Score 554.5; DB 4; Length 1737;
Best Local Similarity 21.6%; Pred. No. 7.7e-38;
Matches 291; Conservative 193; Mismatches 466; Indels 397; Gaps 56;

Qy 18 PRGPSSAGADTN-----STPS-----GSSSGSTGETEYAAAR-----EK 49
Db 381 PYDPEDPGQETINSMSPISQAPDIGRKLERLEDKKNKTLGDLVREAEIFNKRETPER 440
Qy 50 TERAERTIQTGSDR-----GLTAPRAGGDTIOGATN 80
Db 441 EERIRTEBEKEARRTEDEQEKERDRRRHREMSKLLATVVSQKODRGGERRRSQLD 500
```

Db	1546	GOVERNRTKETLTLTATGSRDWWLLPLALYRARNTPGPHGLTPYEILYGAPPLV	1605
QY	944	-----INMITTEQEI-FLQ-----KNSKLKDRVYFREGRQ-----LWKG	980
Db	1606	NFPDPMTRVTSPLQALHQAALYLVQHEWVRPLAAAYQEQLDPRVPVPHYRGVDIVVVR	1665
QY	981	PG-----ELLWKGEAVLVKVGTDIKI	1002
Db	1666	RHQTKNLEPRWKGPYTLVTLTPTALKV	1692

Search completed: September 25, 2003, 12:04:41
Job time : 33 secs

THIS PAGE BLANK (USPTO)

!!AA_MULTIPLE_ALIGNMENT 1.0
PileUp of: *.swissprot

Symbol comparison table: GenRunData:biosum62.cmp CompCheck: 1102

GapWeight: 8
GapLengthWeight: 2

sp.msrf MSF: 1165 Type: P September 25, 2003 10:17 Check: 5858

Name: pol_hv2d1	Len: 1165	Check: 9753	Weight: 1.00
Name: pol_hv2g1	Len: 1165	Check: 9239	Weight: 1.00
Name: pol_hv2be	Len: 1165	Check: 307	Weight: 1.00
Name: pol_hv2ca	Len: 1165	Check: 9370	Weight: 1.00
Name: pol_hv2ro	Len: 1165	Check: 4767	Weight: 1.00
Name: pol_hv2nz	Len: 1165	Check: 8741	Weight: 1.00
Name: pol_hv2kr	Len: 1165	Check: 8455	Weight: 1.00
Name: pol_hv2st	Len: 1165	Check: 9818	Weight: 1.00
Name: pol_hv2sb	Len: 1165	Check: 2098	Weight: 1.00
Name: pol_hv2s4	Len: 1165	Check: 6726	Weight: 1.00
Name: pol_hv2d2	Len: 1165	Check: 6584	Weight: 1.00

//

1	VLELWKGGLT	GETVSTQKT	GLLEVMQVRT	HGKLPKGTG	RFRDQSGTK	50
pol_hv2d1	-----	-----	-----	-----	-----	
pol_hv2g1	-----	-----	-----	-----	-----	
pol_hv2be	-----	-----	-----	-----	-----	
pol_hv2ca	-----	-----	-----	-----	-----	
pol_hv2ro	-----	-----	-----	-----	-----	
pol_hv2nz	-----	-----	-----	-----	-----	
pol_hv2kr	-----	-----	-----	-----	-----	
pol_hv2st	-----	-----	-----	-----	-----	
pol_hv2sb	-----	-----	-----	-----	-----	
pol_hv2s4	-----	-----	-----	-----	-----	
pol_hv2d2	-----	-----	-----	-----	-----	
51	AAQPLPRGGS	SSGADTNSPT	NRSSSG...	P	VGEIYAAREK	AERAEGETIQ
pol_hv2d1	-----	-----	-----	-----	-----	100
pol_hv2g1	-----	-----	-----	-----	-----	
pol_hv2be	-----	-----	-----	-----	-----	
pol_hv2ca	-----	-----	-----	-----	-----	
pol_hv2ro	-----	-----	-----	-----	-----	
pol_hv2nz	-----	-----	-----	-----	-----	
pol_hv2kr	-----	-----	-----	-----	-----	
pol_hv2st	-----	-----	-----	-----	-----	
pol_hv2sb	-----	-----	-----	-----	-----	
pol_hv2s4	-----	-----	-----	-----	-----	
pol_hv2d2	-----	-----	-----	-----	-----	
101	GGDGLTAPR	AGRDAPO	RG	DRGLATPOFS	LWKRPPVTAFA	IEDQPVLEVLL
pol_hv2d1	-----	-----	-----	-----	-----	150
pol_hv2g1	-----	-----	-----	-----	-----	
pol_hv2be	-----	-----	-----	-----	-----	
pol_hv2ca	-----	-----	-----	-----	-----	
pol_hv2ro	-----	-----	-----	-----	-----	
pol_hv2nz	-----	-----	-----	-----	-----	
pol_hv2kr	-----	-----	-----	-----	-----	
pol_hv2st	-----	-----	-----	-----	-----	
pol_hv2sb	-----	-----	-----	-----	-----	
pol_hv2s4	-----	-----	-----	-----	-----	
pol_hv2d2	-----	-----	-----	-----	-----	
151	GTGADDSIVA	GIELGDNVTP	KIVGGIGGFI	NTRYKKNVEI	KVLNKKVRAT	200
pol_hv2d1	-----	-----	-----	-----	-----	
pol_hv2g1	-----	-----	-----	-----	-----	
pol_hv2be	-----	-----	-----	-----	-----	
pol_hv2ca	-----	-----	-----	-----	-----	
pol_hv2ro	-----	-----	-----	-----	-----	
pol_hv2nz	-----	-----	-----	-----	-----	
pol_hv2kr	-----	-----	-----	-----	-----	
pol_hv2st	-----	-----	-----	-----	-----	
pol_hv2sb	-----	-----	-----	-----	-----	

pol_hv2Ro

90

IMTGTPTINI
FGRNII-LJTAG
MSLNLPUAKN

pol_hv2s4	DTGADDSIVA	GIELGPNYTP	KIVGGIGGFI	NTRYKDVKI	KVLGKVIKST	
pol_hv2d2	DTGVDDSDIVA	GIELGNSYTP	KIVGGIGGFI	NTRYKDVKI	EVVGRVRAT	
201	..	90				
pol_hv2d1	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKL	DPIKVTLPK	KDGPRLKQWP	250
pol_hv2g1	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKI	EPIKVTLPK	KDGPRLRQWP	
pol_hv2be	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKI	EPIKVTLPK	KDGPRLRQWP	
pol_hv2ca	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKI	EPIKVTLPK	KDGPRLRQWP	
pol_hv2ro	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKI	EPIKVTLPK	KDGPRLRQWP	
pol_hv2nz	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKI	EPIKVTLPK	KDGPRLRQWP	
pol_hv2kr	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKI	EPIKVTLPK	KDGPRLRQWP	
pol_hv2st	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKI	EPIKVTLPK	KDGPRLRQWP	
pol_hv2sb	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKI	EPIKVTLPK	KDGPRLRQWP	
pol_hv2s4	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKI	EPIKVTLPK	KDGPRLRQWP	
pol_hv2d2	IMTGDTPINI	FGNIIATLIG	MTLNPVAVK	EPVAVELKPG	KDGPKLRQWP	
251						300
pol_hv2d1	LTRKIEALK	EICEKMEREG	OLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
pol_hv2g1	LTRKIEALR	EICEKMEREG	OLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
pol_hv2be	LTRKIEALR	EICEKMEREG	OLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
pol_hv2ca	LTRKIEALK	EICEKMEREG	OLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
pol_hv2ro	LTRKIEALK	EICEKMEREG	OLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
pol_hv2nz	LTRKIEALK	EICEKMEREG	OLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
pol_hv2kr	LTRKIEALK	EICEKMEREG	OLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
pol_hv2st	LTRKIEALK	EICEKMEREG	OLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
pol_hv2sb	LTRKIEALR	EICEKMEREG	OLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
pol_hv2s4	LSREKIIALR	EICEKMEREG	OLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
pol_hv2d2	LSREKILALK	EICEKMEREG	OLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
301						350
pol_hv2d1	DFRELNRVQT	DFTETQLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
pol_hv2g1	DFRELNRVQT	DFTETQLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
pol_hv2be	DFRELNRVQT	DFTETQLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
pol_hv2ca	DFRELNRVQT	DFTETQLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
pol_hv2ro	DFRELNRVQT	DFTETQLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
pol_hv2nz	DFRELNRVQT	DFTETQLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
pol_hv2kr	DFRELNRVQT	DFTETQLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
pol_hv2st	DFRELNRVQT	DFTETQLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
pol_hv2sb	DFRELNRVQT	DFTETQLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
pol_hv2s4	DFRELNRVQT	DFTETQLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
pol_hv2d2	DFRELNRVQT	DFTETVNWVFP	.TRQVAEKRR	ITVLVDGDAY	FSIPLDNPFR	
351						400
pol_hv2d1	QYTAFTLPSP	NNAEPGRVY	YKVLPOGWKG	SPAIFOPMMR	QILEPFRKAN	
pol_hv2g1	QYTAFTLPSP	NNAEPGRVY	YKVLPOGWKG	SPAIFOPMMR	QILEPFRKAN	
pol_hv2be	QYTAFTLPSP	NNAEPGRVY	YKVLPOGWKG	SPAIFOPMMR	QILEPFRKAN	
pol_hv2ca	QYTAFTLPSP	NNAEPGRVY	YKVLPOGWKG	SPAIFOPMMR	QILEPFRKAN	
pol_hv2ro	QYTAFTLPSP	NNAEPGRVY	YKVLPOGWKG	SPAIFOPMMR	QILEPFRKAN	
pol_hv2nz	QYTAFTLPSP	NNAEPGRVY	YKVLPOGWKG	SPAIFOPMMR	QILEPFRKAN	
pol_hv2kr	QYTAFTLPSP	NNAEPGRVY	YKVLPOGWKG	SPAIFOPMMR	QILEPFRKAN	
pol_hv2st	QYTAFTLPSP	NNAEPGRVY	YKVLPOGWKG	SPAIFOPMMR	QILEPFRKAN	
pol_hv2sb	QYTAFTLPSP	NNAEPGRVY	YKVLPOGWKG	SPAIFOPMMR	QILEPFRKAN	
pol_hv2s4	QYTAFTLPSP	NNAEPGRVY	YKVLPOGWKG	SPAIFOPMMR	QILEPFRKAN	
pol_hv2d2	QYTAFTLPSP	NNAEPGRVY	YKVLPOGWKG	SPAIFOPMMR	QILEPFRKAN	
401						450
pol_hv2d1	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQL	KELLNGLGFS	TPDEKFQKDP	
pol_hv2g1	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQL	KELLNGLGFS	TPDEKFQKDP	
pol_hv2be	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQL	KELLNGLGFS	TPDEKFQKDP	
pol_hv2ca	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQL	KELLNGLGFS	TPDEKFQKDP	
pol_hv2ro	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQL	KELLNGLGFS	TPDEKFQKDP	
pol_hv2nz	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQL	KELLNGLGFS	TPDEKFQKDP	
pol_hv2kr	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQL	KELLNGLGFS	TPDEKFQKDP	
pol_hv2st	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQL	KELLNGLGFS	TPDEKFQKDP	
pol_hv2sb	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQL	KELLNGLGFS	TPDEKFQKDP	
pol_hv2s4	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQL	KELLNGLGFS	TPDEKFQKDP	
pol_hv2d2	PDVILIIQYMD	DILIASDRSD	LEHDKVVSQ	KELLNDMGFS	TPDEKFQKDP	
451						500
pol_hv2d1	PFOWMGYELW	PTKWKLOKIQ	LPOKEIWTVN	DIQKLVGVLN	WAAQIYPGIK	
pol_hv2g1	PFOWMGYELW	PTKWKLOKIQ	LPOKEIWTVN	DIQKLVGVLN	WAAQIYPGIK	
pol_hv2be	PFOWMGYELW	PTKWKLOKIQ	LPOKEIWTVN	DIQKLVGVLN	WAAQIYPGIK	
pol_hv2ca	PFOWMGYELW	PTKWKLOKIQ	LPOKEIWTVN	DIQKLVGVLN	WAAQIYPGIK	

751	pol_hv2d1	KREAVYVAVY	PAHKGIGGNQ	EVDBHLSQGI	QVLFLEKTE	PAQBEHEKYYH	801
	pol_hv2g1	KREAVYVAVY	PAHKGIGGNQ	EVDBHLSQGI	QVLFLEKTE	PAQBEHEKYYH	
	pol_hv2be	KREAVYVAVY	PAHKGIGGNQ	EVDBHLSQGI	QVLFLEKTE	PAQBEHEKYYH	
	pol_hv2ca	KREAVYVAVY	PAHKGIGGNQ	EVDBHLSQGI	QVLFLEKTE	PAQBEHEKYYH	
	pol_hv2ro	KREAVYVAVY	PAHKGIGGNQ	EVDBHLSQGI	QVLFLEKTE	PAQBEHEKYYH	
	pol_hv2nz	KREAVYVAVY	PAHKGIGGNQ	EVDBHLSQGI	QVLFLEKTE	PAQBEHEKYYH	
	pol_hv2kr	KREAVYVAVY	PAHKGIGGNQ	EVDBHLSQGI	QVLFLEKTE	PAQBEHEKYYH	
	pol_hv2st	KREAVYVAVY	PAHKGIGGNQ	EVDBHLSQGI	QVLFLEKTE	PAQBEHEKYYH	
	pol_hv2sb	KREAVYVAVY	PAHKGIGGNQ	EVDBHLSQGI	QVLFLEKTE	PAQBEHEKYYH	
	pol_hv2s4	KREAVYVAVY	PAHKGIGGNQ	EVDBHLSQGI	QVLFLEKTE	PAQBEHEKYYH	
pol_hv2d2	KREAVYVGVY	PAHKGLGGNQ	EVDBHLSQGI	QVLFLEKTE	PAQBEHEKYYH		
801	pol_hv2d1	SNIKELTHRF	GIPQLVARQI	VNTCAOQOQK	GEAIGHGOVNA	EIGVWQMDCT	850
	pol_hv2g1	SNIKELTHRF	GIPQLVARQI	VNTCAOQOQK	GEAIGHGOVNA	EIGVWQMDCT	
	pol_hv2be	SNIKELTHRF	GIPQLVARQI	VNTCAOQOQK	GEAIGHGOVNA	EIGVWQMDCT	
	pol_hv2ca	SNIKELCHRF	DIPQLVARQI	VNTCAOQOQK	GEAIGHGOVNA	ELGTWQMDCT	
	pol_hv2ro	SNIKELCHRF	DIPQLVARQI	VNTCAOQOQK	GEAIGHGOVNA	ELGTWQMDCT	
	pol_hv2nz	SNIKELSHRF	GIPKLVARQI	VNTCAHVOQK	GEAIGHGOVNA	ELGTWQMDCT	
	pol_hv2kr	SNIKELSHRF	GIPKLVARQI	VNTCAHVOQK	GEAIGHGOVNA	ELGTWQMDCT	
	pol_hv2st	SNIKELSHRF	GIPKLVARQI	VNTCAHVOQK	GEAIGHGOVNA	ELGTWQMDCT	
	pol_hv2sb	SNIKELAHRF	GLPNLVARQI	VNTCAOQOQK	GEAIGHGOVNA	ELGTWQMDCT	
	pol_hv2s4	SNIKELVLFHF	GLPNLVARQI	VNTCAOQOQK	GEAIGHGOVNA	ELGTWQMDCT	
pol_hv2d2	GNVKELVHF	GIPQLVARQI	VNSCDKCOQK	GEAIGHGOVNA	DLGTWQMDCT		
851	pol_hv2d1	HLEGKIIIVA	VHVASGFTEA	EVIPQESGRQ	TALFLKLKAS	RWPITHLHTD	900
	pol_hv2g1	HLEGKIIIVA	VHVASGFTEA	EVIPQESGRQ	TALFLKLKAS	RWPITHLHTD	
	pol_hv2be	HLEGKIIIVA	VHVASGFTEA	EVIPQESGRQ	TALFLKLKAS	RWPITHLHTD	
	pol_hv2ca	HLEGKIIIVA	VHVASGFTEA	EVIPQESGRQ	TALFLKLKAS	RWPITHLHTD	
	pol_hv2ro	HLEGKIIIVA	VHVASGFTEA	EVIPQESGRQ	TALFLKLKAS	RWPITHLHTD	
	pol_hv2nz	HLEGKIIIVA	VHVASGFTEA	EVIPQESGRQ	TALFLKLKAS	RWPITHLHTD	
	pol_hv2kr	HLEGKIIIVA	VHVASGFTEA	EVIPQESGRQ	TALFLKLKAS	RWPITHLHTD	
	pol_hv2st	HLEGKIIIVA	VHVASGFTEA	EVIPQESGRQ	TALFLKLKAS	RWPITHLHTD	
	pol_hv2sb	HLEGKIIIVA	VHVASGFTEA	EVIPQESGRQ	TALFLKLKAS	RWPITHLHTD	
	pol_hv2s4	HLEGKIIIVA	VHVASGFTEA	EVIPQETGRQ	TALFLKLKAS	RWPITHLHTD	
pol_hv2d2	HLEGKIIIVA	VHVASGFTEA	EVIPQETGRQ	TALFLKLKAS	RWPITHLHTD		
901	pol_hv2d1	NGPNFTSQEV	KVAVWVIGIE	QSGFVVPYNPQ	SQGVVEAMNH	HLKNQISIRIR	950
	pol_hv2g1	NGSNFTSQEV	KVAVWVIGIE	QSGFVVPYNPQ	SQGVVEAMNH	HLKNQISIRIR	
	pol_hv2be	NGPNFTSQEV	KVAVWVIGIE	QSGFVVPYNPQ	SQGVVEAMNH	HLKNQISIRIR	
	pol_hv2ca	NGANFTSQEV	KVAVWVIGIE	QSGFVVPYNPQ	SQGVVEAMNH	HLKNQISIRIR	
	pol_hv2ro	NGANFTSQEV	KVAVWVIGIE	QSGFVVPYNPQ	SQGVVEAMNH	HLKNQISIRIR	
	pol_hv2nz	NGANFTSQEV	KVAVWVIGIE	QSGFVVPYNPQ	SQGVVEAMNH	HLKNQISIRIR	
	pol_hv2kr	NGANFTSQEV	KVAVWVIGIE	QSGFVVPYNPQ	SQGVVEAMNH	HLKNQISIRIR	
	pol_hv2st	NGANFTSQEV	KVAVWVIGIE	QSGFVVPYNPQ	SQGVVEAMNH	HLKNQISIRIR	
	pol_hv2sb	NGANFTSQEV	KVAVWVIGIE	QSGFVVPYNPQ	SQGVVEAMNH	HLKNQISIRIR	
	pol_hv2s4	NGANFTSQEV	KVAVWVIGIE	QSGFVVPYNPQ	SQGVVEAMNH	HLKNQISIRIR	
pol_hv2d2	NGANFTSPSV	KVAVWVIGIE	QTFGVVPYNPQ	SQGVVEAMNH	HLKNQIDRLR		
951	pol_hv2d1	EQANTTETIV	LMVAVHCNFK	RGGIGDMTP	ABRLNMNMT	EOETQFLORK	1000
	pol_hv2g1	EQANTTETIV	LMVAVHCNFK	RGGIGDMTP	ABRLNMNMT	EOETQFLORK	
	pol_hv2be	EQANTTETIV	LMVAVHCNFK	RGGIGDMTP	ABRLNMNMT	EOETQFLORK	
	pol_hv2ca	EQANTTETIV	LMVAVHCNFK	RGGIGDMTP	ABRLNMNMT	EOETQFLORK	
	pol_hv2ro	EQANTTETIV	LMVAVHCNFK	RGGIGDMTP	ABRLNMNMT	EOETQFLORK	
	pol_hv2nz	EQANTTETIV	LMVAVHCNFK	RGGIGDMTP	ABRLNMNMT	EOETQFLORK	
	pol_hv2kr	EQANTTETIV	LMVAVHCNFK	RGGIGDMTP	ABRLNMNMT	EOETQFLORK	
	pol_hv2st	EQANTTETIV	LMVAVHCNFK	RGGIGDMTP	ABRLNMNMT	EOETQFLORK	
	pol_hv2sb	EQANTTETIV	LMVAVHCNFK	RGGIGDMTP	ABRLNMNMT	EOETQFLORK	
	pol_hv2s4	EQANSTETIV	LMVAVHCNFK	RGGIGDMTP	ABRLNMNMT	EOETQFLORK	
pol_hv2d2	EQANSTETIV	LMVAVHCNFK	RGGIGDMTP	ABRLNMNMT	EOETQFLORK		
1001	pol_hv2d1	NSNFKFQVY	YREGDQLWK	GPCELLWKGD	GAVIVKVGAD	IKVYPRRKAK	1050
	pol_hv2g1	NSNFKNFQY	YREGDQLWK	GPCELLWKGD	GAVIVKVGAD	IKVYPRRKAK	
	pol_hv2be	NSNFKNFQY	YREGDQLWK	GPCELLWKGD	GAVIVKVGAD	IKVYPRRKAK	
	pol_hv2ca	NSKLKDFRY	FREGDQL				

```
pol_hv2st  NSKLQNRVY  FREGRDQLWK  GPGELLWKGD  GAVIVKVGAD  IKIIPRRKAK
pol_hv2sb  NSKLKNRVY  FREGRDQLWK  GPGELLWKGD  GAVIVKVGTD  IKVIPRRKAK
pol_si4    NSKFKNRVY  FREGRDQLWK  GPGELLWKGE  GAVILKVGTE  IKVVPRRKAK
pol_hv2d2  NLKFNQFVY  YREGRDQLWK  GPGELLWKGE  GAVILKVGTE  IKVVPRRKAK

1051
pol_hv2d1  IIRDYGRQE  LDSSHLEGA  R.EDGEVA--  -----  1100
pol_hv2g1  IIRDYGRQE  LD.  SSHLEGA  REEDGEVA--  -----
pol_hv2be  IIRDYGRQE  LDSSHLEGA  R.EDGEVACP  CQVPEIONKR  PRGGALCSPP
pol_hv2ca  IIRDYGRQE  LDSSHLEGA  R.EDGEVA--  -----
pol_hv2ro  IIRDYGRQE  MDSSGHLEGA  R.EDGEVA--  -----
pol_hv2nz  IIRDYGRQE  MDSSGHLEGA  R.EDGEVA--  -----
pol_hv2kr  IIRDYGRQE  MDSSGHLEGA  R.EDGEVA--  -----
pol_hv2st  IIRDYGRQE  MDSSGHLEGA  R.EDGEVA--  -----
pol_hv2sb  IIRDYGRQE  MDSSGHLEGA  R.EDGEVA--  -----
pol_si4    IIRDYGRQE  MDSSGHLEGA  R.EDGEVA--  -----
pol_hv2d2  IIRHYGGKG  LDSCADMET  R.QAREMAQS  D-----

1101
pol_hv2d1  -----  -----  -----  -----  1150
pol_hv2g1  -----  -----  -----  -----
pol_hv2be  QGGMGWVDLQ  OGNIPTRKK  SSRNTGILEP  NTRKRMALLS  CSKINLVYRK
pol_hv2ca  -----  -----  -----  -----
pol_hv2ro  -----  -----  -----  -----
pol_hv2nz  -----  -----  -----  -----
pol_hv2kr  -----  -----  -----  -----
pol_hv2st  -----  -----  -----  -----
pol_hv2sb  -----  -----  -----  -----
pol_si4    -----  -----  -----  -----
pol_hv2d2  -----  -----  -----  -----

1151
pol_hv2d1  -----  1165
pol_hv2g1  -----
pol_hv2be  VLDRCYPRLC  RHPNT
pol_hv2ca  -----
pol_hv2ro  -----
pol_hv2nz  -----
pol_hv2kr  -----
pol_hv2st  -----
pol_hv2sb  -----
pol_si4    -----
pol_hv2d2  -----
```

THIS PAGE BLANK (USPTO)

!!AA_MULTIPLE_ALIGNMENT 1.0

Pileup of: *.sp_rvirus

Symbol comparison table: GenRunData:blosum62.cmp CompCheck: 1102

GapWeight: 8

GapLengthWeight: 2

pileup.msf MSF: 1645 Type: P September 25, 2003 10:13 Check: 2437

Name: q76053 Len: 1645 Check: 4352 Weight: 1.00
Name: q90066 Len: 1645 Check: 2138 Weight: 1.00
Name: q77162 Len: 1645 Check: 4711 Weight: 1.00
Name: q77249 Len: 1645 Check: 3717 Weight: 1.00
Name: q73194 Len: 1645 Check: 7255 Weight: 1.00
Name: q78654 Len: 1645 Check: 3402 Weight: 1.00
Name: q76630 Len: 1645 Check: 4584 Weight: 1.00
Name: q9ytu0 Len: 1645 Check: 9112 Weight: 1.00
Name: q78655 Len: 1645 Check: 3199 Weight: 1.00
Name: q89928 Len: 1645 Check: 4930 Weight: 1.00
Name: q9ib19 Len: 1645 Check: 9116 Weight: 1.00
Name: q9dsk7 Len: 1645 Check: 5921 Weight: 1.00

//

1
q76053
q90066
q77162
q77249
q73194
q78654
q76630
q9ytu0
q78655
q89928
q9ib19
q9dsk7
TSTCYMRHCN TSVIKESCDK HYWDAMKPRY CAPPGFALLR CNDNTYSGFE
50

51
q76053
q90066
q77162
q77249
q73194
q78654
q76630
q9ytu0
q78655
q89928
q9ib19
q9dsk7
PNC SKVVATT CTRMETQTS TWFGPNSRA ENRTYIYVHS KDNRTIISLN
100

101
q76053
q90066
q77162
q77249
q73194
q78654
q76630
q9ytu0
q78655
q89928
q9ib19
q9dsk7
SFYNLTWHCK RPKGNTVVP I TLMGLVPHS QPINKRPRQA KCWFGGNWTG
150

151
q76053
q90066
q77162
q77249
q73194
q78654
q76630
q9ytu0
q78655
q89928
q9ib19
q9dsk7
200

q78654
q76630
q9ytu0
q78655
q89928
q9ib19
q9dsk7
AMQEVKOTLA KTSQVORNHD TKKINFARPG MGSDEPVVYM WTNCRGEFLY
201
q76053
q90066
q77162
q77249
q73194
q78654
q76630
q9ytu0
q78655
q89928
q9ib19
q9dsk7
CNMTWFLNWV ENRTNTTOHN YAPCHIKQII NAWHKVGKYY YLPPREGELT
250

251
q76053
q90066
q77162
q77249
q73194
q78654
q76630
q9ytu0
q78655
q89928
q9ib19
q9dsk7
CNSTVTSIIA NIDTGANQIN ITFSAEVSDL YRLEGDYKL IEITPIGFAP
300

301
q76053
q90066
q77162
q77249
q73194
q78654
q76630
q9ytu0
q78655
q89928
q9ib19
q9dsk7
CNSTVTSIIA NIDTGANQIN ITFSAEVSDL YRLEGDYKL IEITPIGFAP
350

351
q76053
q90066
q77162
q77249
q73194
q78654
q76630
q9ytu0
q78655
q89928
q9ib19
q9dsk7
TNVKRYSSTP VRNKRGVFVL GFLGFLATAG AAMGTASLTL SAQSRTLLAG
400

401
q76053
q90066
q77162
q77249
q73194
q78654
q76630
q9ytu0
q78655
q89928
q9ib19
q9dsk7
IVQQOQOQLLD VVKRQOELLR LTVWGTKNLQ ARVTAIEKYL KHAQALNSMG
450

451
q76053
q90066
q77162
q77249
q73194
q78654
q76630
q9ytu0
q78655
q89928
q9ib19
q9dsk7
450

[illegible]

q9dsk7	701	751	801	851	901	950	1000
q76053	VWTAIEGQP	VEVLDTGAD	DSIVAGIELG	NNYSPIKIVGG	NNYSPIKIVGG	IGGFINTKEY	750
q90066	VWTAIEGQP	VEVLDTGAD	DSIVAGIELG	NNYSPIKIVGG	NNYSPIKIVGG	IGGFINTKEY	
q77162	VWTAIEGQP	VEVLDTGAD	DSIVAGIELG	DNTPKIVGG	DNTPKIVGG	IGGFINTKEY	
q77249	VWTAIEGQP	VEVLDTGAD	DSIVAGIELG	DNTPKIVGG	DNTPKIVGG	IGGFINTKEY	
q73194	VWTAHIEGQP	VEVLDTGAD	DSIVAGIELG	SNYSPIKIVGG	SNYSPIKIVGG	IGGFINTKEY	
q78654	VWTAHIEGQP	VEVLDTGAD	DSIVAGIELG	CDXSPKIVGG	CDXSPKIVGG	IGGFINTKEY	
q76630	VWTAHIEGQP	VEVLDTGAD	DSIVAGIELG	SNYSPIKIVGG	SNYSPIKIVGG	IGGFINTKEY	
q9ytcu0	VWTAIEGQP	VEVLDTGAD	DSIVAGIELG	SNYTPKIVGG	SNYTPKIVGG	IGGFINTKEY	
q78655	VWTAHIEGQP	VEVLDTGAD	DSIVAGIELG	SNYTPKIVGG	SNYTPKIVGG	IGGFINTKEY	
q89928	VWTAHIEGQP	VEVLDTGAD	DSIVAGIELG	SNYTPKIVGG	SNYTPKIVGG	IGGFINTKEY	
q91b19	VWTAIESQP	VEVLDTGAD	DSIVAGIELG	SNYSPIKIVGG	SNYSPIKIVGG	IGGFINTKEY	
q9dsk7							
q76053	KNVEIEVLNK	VRVATIMTGD	TPNIFGRNI	L7ALGMSLNL	L7ALGMSLNL	800	
q90066	KNVEIEVLNK	VRVATIMTGD	TPNIFGRNI	L7ALGMSLNL	L7ALGMSLNL		
q77162	KNVEIEVLNK	VRVATIMTGD	TPNIFGRNI	L7ALGMSLNL	L7ALGMSLNL		
q77249	KNVEIEVLNK	VRVATIMTGD	TPNIFGRNI	L7ALGMSLNL	L7ALGMSLNL		
q73194	KNVEIEVLNK	VRVATIMTGD	TPNIFGRNI	L7ALGMSLNL	L7ALGMSLNL	PVAKTEPIKI	
q78654	KNVEIEVLGR	VRVATIMTGD	TPNIFGRNV	L7ALGMSLNL	L7ALGMSLNL		
q76630	KNVEIEVLNK	VRVATIMTGD	TPNIFGRNI	L7ALGMSLNL	L7ALGMSLNL	PVAKTEPIKI	
q9ytcu0	EDVEIELNK	VRVATIMTGD	TPNIFGRNI	L7ALGMSLNL	L7ALGMSLNL	PVAKTEPIEV	
q78655	KNVKEILLNK	VRVATIMTGD	TPNIFGRNI	L7ALGMSLNL	L7ALGMSLNL		
q89928	KNVEIEVVGK	VRVATIMTGD	TPNIFGRNI	LNSLGMTLNF	LNSLGMTLNF	PVARTPEVKV	
q91b19	KNVEIKVVGK	VRVATIMTGD	TPNIFGRNI	LVKLGMSLNF	LVKLGMSLNF	PLAKVEPVKV	
q9dsk7							
q76053	MLKPGKDGPR	LQWPLTKK	IEALKEICEK	MEKEGOLEEA	MEKEGOLEEA	PPTNPYNTPT	850
q90066	MLKPGKDGPR	LQWPLTKK	IEALKEICEK	MEKEGOLEEA	MEKEGOLEEA	PPTNPYNTPT	
q77162	MLKPGKDGPR	LQWPLTKK	IEALKEICEK	MEKEGOLEEA	MEKEGOLEEA	PPTNPYNTPT	
q77249	MLKPGKDGPR	LQWPLTKK	IEALKEICEK	MEKEGOLEEA	MEKEGOLEEA	PPTNPYNTPT	
q73194	MLKPGKDGPR	LQWPLTKK	IEALKEICEK	MEKEGOLEEA	MEKEGOLEEA	PPTNPYNTPT	
q78654	MLKPGKDGPR	LQWPLTKK	IEALKEICEK	MEKEGOLEEA	MEKEGOLEEA	PPTNPYNTPT	
q76630	MLKPGKDGPR	LQWPLTKK	IEALKEICEK	MEKEGOLEEA	MEKEGOLEEA	PPTNPYNTPT	
q9ytcu0	MLKPGKDGPR	LQWPLTKK	IEALKEICEK	MEKEGOLEEA	MEKEGOLEEA	PPTNPYNTPT	
q78655	MLKPGKDGPR	LQWPLTKK	IEALKEICEK	MEKEGOLEEA	MEKEGOLEEA	PPTNPYNTPT	
q89928	MLKPGKDGPR	LQWPLTKK	IEALKEICEK	MEKEGOLEEA	MEKEGOLEEA	PPTNPYNTPT	
q91b19	MLKPGKDGPR	LQWPLTKK	IEALKEICEK	MEKEGOLEEA	MEKEGOLEEA	PPTNPYNTPT	
q9dsk7							
q76053	OLGIPHPAGL	AKKRITVLD	OLGIPHPAGL	AKKRITVLD	OLGIPHPAGL	900	
q90066	OLGIPHPAGL	AKKRITVLD	OLGIPHPAGL	AKKRITVLD	OLGIPHPAGL		
q77162	OLGIPHPAGL	AKKRITVLD	OLGIPHPAGL	AKKRITVLD	OLGIPHPAGL		
q77249	OLGIPHPAGL	AKKRITVLD	OLGIPHPAGL	AKKRITVLD	OLGIPHPAGL		
q73194	OLGIPHPAGL	AKKRITVLD	OLGIPHPAGL	AKKRITVLD	OLGIPHPAGL		
q78654	OLGIPHPAGL	AKKRITVLD	OLGIPHPAGL	AKKRITVLD	OLGIPHPAGL		
q76630	OLGIPHPAGL	AKKRITVLD	OLGIPHPAGL	AKKRITVLD	OLGIPHPAGL		
q9ytcu0	OLGIPHPAGL	AKKRITVLD	OLGIPHPAGL	AKKRITVLD	OLGIPHPAGL		
q89928	OLGIPHPAGL	AKKRITVLD	OLGIPHPAGL	AKKRITVLD	OLGIPHPAGL		
q91b19	OLGIPHPAGL	AKKRITVLD	OLGIPHPAGL	AKKRITVLD	OLGIPHPAGL		
q9dsk7							
q76053	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	950	
q90066	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF		
q77162	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF		
q77249	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF		
q73194	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF		
q78654	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF		
q76630	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF		
q9ytcu0	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF		
q78655	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF		
q89928	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF		
q91b19	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF	QWKGSPAIF		
q9dsk7							

q76630	LKLSRWPIT	HLHTDNGANF	TQOEVKMAV	WIGIEQSEGV	PYNQSQGVV
q9ylu0	LKLSRWPIT	HLHTDNGVNF	TQOEVKMAV	WVIGIEQSEGV	PYNQSQGVV
q78655	-----	-----	-----	-----	-----
q89928	LKLSRWPIT	HLHTDNGANF	TQODVKMAV	WIGIEQTEGV	PYNPESQGVV
q91b19	LKLAGRWPIT	HLHTDNGANF	TQOEVKMAV	WVIGIEQAEV	PYNQSQGVV
q9dsk7	-----	-----	-----	-----	-----
1501	-----	-----	-----	-----	1550
q76053	-----	-----	-----	-----	-----
q90066	-----	-----	-----	-----	-----
q77162	-----	-----	-----	-----	-----
q77249	-----	-----	-----	-----	-----
o73194	EAMNHHLKQ	ISRIREQANT	VETIVLMAVH	CMNFKRRGGI	GMTTPSERLI
q78654	-----	-----	-----	-----	-----
q76630	EAMNHHLKQ	ISRIREQANT	VETIVLMAVH	CMNFKRRGGI	GMTTPAERLI
q9ylu0	EAMNHHLKQ	ISRIREQANT	VETIVLMAVH	CMNFKRRGGI	GMTTPAERLI
q78655	-----	-----	-----	-----	-----
q89928	EAMNHHLKQ	IDRIREQAVS	IETIVLMATH	CMNFKRRGGI	GMTTPAERIV
q91b19	ESMNHHLKQ	IDKIRDOANS	IETIVLMAVH	CMNFKRRGGI	GMTTPAERII
q9dsk7	-----	-----	-----	-----	-----
1551	-----	-----	-----	-----	1600
q76053	-----	-----	-----	-----	-----
q90066	-----	-----	-----	-----	-----
q77162	-----	-----	-----	-----	-----
q77249	-----	-----	-----	-----	-----
o73194	NMITTEQEIQ	FLQAKNSKLK	NFRVYFREG	DQLMKGPGEI	LWKGDGAVIV
q78654	-----	-----	-----	-----	-----
q76630	NMTAEQEIQ	FLQAKNSKLQ	NFRVYFREG	DQLMKGPGEI	LWKGDGAVIV
q9ylu0	NMISTEQEIQ	FLQTKNLKFK	NFRVYFREG	DQLMKGPGEI	LWKGDGAVIV
q78655	-----	-----	-----	-----	-----
q89928	NMITTEQEIQ	FLQTKNLKFK	NFRVYFREG	DQLMKGPGEI	LWKGEQAVII
q91b19	NMITTEQEIQ	FOQTKNSKFK	NFRVYFREG	DQLMKGPGEI	LWKGEQAVII
q9dsk7	-----	-----	-----	-----	-----
1601	-----	-----	-----	-----	1645
q76053	-----	-----	-----	-----	-----
q90066	-----	-----	-----	-----	-----
q77162	-----	-----	-----	-----	-----
q77249	-----	-----	-----	-----	-----
o73194	KVGTDIKIIP	RRKAKIIRDY	GGQELDSSS	HLEGARENGE	VA---
q78654	-----	-----	-----	-----	-----
q76630	KVGADIKIIP	RRKAKIKDY	GGQEMDSCS	NLEGAREDEGE	VA---
q9ylu0	KVGTDIKVVP	RRKAKIIRDY	GGQELDSCP	HLEGAREDEGE	VA---
q78655	-----	-----	-----	-----	-----
q89928	KVGTEIKVIP	RRKAKIIRNY	GGGKELDCSA	DVEDTMOARE	VAQSN
q91b19	KVGTEIKVVP	RRKAKIKDY	GGGKEVDSST	NMEDTROTGE	VA---
q9dsk7	-----	-----	-----	-----	-----